The "Social" in Social Science:
The Implications of Social Networks Theory for Political Economy and Political Methodology

Neelanjan Sircar

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2014

ABSTRACT

The "Social" in Social Science

Neelanjan Sircar

This dissertation focuses on the consequences of considering social networks in standard frameworks in political economy and political methodology. The first paper introduces a game theoretic model where an allocator may distribute benefits over a social network, but the units in the network may extract rents from the allocator. This amendment to a classic allocation game generates unique predictions. Units can use their social position to extract rents and corruption persists in equilibrium. This has major implications for the provision of local public goods. The second paper derives a general statistical framework for causal identification in randomized experiments in the presence of spillovers. The paper addresses a major question in the analysis of randomized experiments and develops a framework that can be readily applied by practitioners. The third paper analyzes the impact of kinship networks on political preference formation in rural India. It is shown that kinship networks help to pool information, generate political discussion and provide explicit coordination of political behavior. This provides an account of how social structure functions in democratic developing country contexts.

Table of Contents

List of Figures

List of Tables

Acknowledgments

The number of people who helped me over the course of my dissertation is so great that it is simply not feasible to name everyone. I am deeply indebted to all of those who helped me through my research, in the field, or simply by providing friendship and advice. Even if I have forgotten to acknowledge you, rest assured that I remember you and deeply appreciate what you have done for me.

Don Green and Jake Bowers provided insightful advice on cleaning up and explicating my statistical techniques.

Massimo Morelli has been instrumental in shaping and supporting my work in game theory. It is because of him that I now have the ability to present game-theoretic work in public academic settings. More importantly, he has always been serious about keeping tabs on me, even when I was trying to disappear.

Macartan Humphreys played the role of a jack-of-all-trades throughout this process. The breadth of his knowledge is truly astounding. His careful feedback, combined with his ability to comment on statistics, game theory, and field research, was instrumental in allowing me to complete my dissertation. Above all, I have learned from him the value of being creative and imaginative in carrying out my research.

Philip Oldenburg is the man who first taught me about India. His ideas and tutelage are reflected throughout this entire dissertation. He knew when to ask tough questions of my work and when to be encouraging. I would never have succeeded in carrying out field research in India without him.

To most people, Devesh Kapur is a brilliant academic working on India. But to me, he is a close friend, one whose academic advice and judgment I trust above all else. He took me into his institute, the Center for the Advanced Study of India (CASI) at the University of Pennsylvania, and he is most responsible for my growth as a scholar. In addition to Devesh Kapur, I would like to thank the entire staff at CASI, especially Juliana Di Giustini, whose compassion and support has gotten me through the last stages of this process.

Writing a dissertation is a tough process, and I want to thank all of those who have read my work over the years, or simply been there for me when I needed them. In particular, I want to thank Milan Vaishnav. I took my first academic trip to India with him, and I appreciate the depth and clarity of his advice; most of all, I appreciate his friendship. Peter van der Windt has been an unrelenting supporter, close friend, and co-author; I know we will continue to work together for many years. Mark Schneider has been a close friend, and the person who helped me the most when things weren't going the way I hoped. I would also like to thank Simon Collard-Wexler, Felix Gerlsbeck, Pierce O'Reilly, Francesca Refsum Jensenius, Pavithra Suryanarayan, and Andrej Tusicisny, who have offered friendship and advice over the years.

Finally, I want to thank my family. I want to thank Sumit Sircar, Subrota Sircar, and Indrani Tagore for taking care of me in India when I couldn't take care of myself. I want to thank my brother Indraneel Sircar for blazing the academic trail in the family and making things much easier for me through his advice and encouragement. Finally, I want to thank my parents, Partha and Devjani Sircar, for their unending love and support. Their energy carried me through the last stages of writing my dissertation when my own energy was flagging. My theories on the importance of a kinship network are no doubt influenced by the support and care I received from my own kinship network.

Chapter 1

Introduction

1.1 General Approach

This dissertation combines three disparate papers comprising game theory, the statistical theory of causal inference, and voter behavior in rural India. The tie that binds these papers together is that they each use social networks theory to extend upon major methodological or theoretical approaches in political economy.

As one reads through the dissertation, it will become immediately obvious that it is not intended to be a single, cogent presentation. Rather, the three papers represent separate research agendas. The underlying motivation to write each of these papers was strikingly similar. Each of the papers takes a standard framework in political economy or methodology and demonstrates how things change when one introduces "spillovers" between the units of study. In broadest terms, a spillover occurs any time changing some outcome of interest (e.g., vote preference) in one unit affects the same outcome for another unit in the study. Typically, these spillovers occur through some social relation between the units of study, and the entire collection of such social relations comprise a *social network*. In this dissertation, I will use the term "social network" in a very precise way. It describes a structure containing two sets, one set describing units for study, and one set describing the presence/absence of a social relation between any pair of units in the study. In mathematics, this structure is referred to as a *mathematical graph*.

1.2 The Three Papers

The three papers in this dissertation address three very different areas of substantive and methodological inquiry, each of which underscore the importance of considering social networks.

1.2.1 On Rent Extraction and Efficient Allocation over Social Networks

This paper develops a formal model of allocation and rent-seeking over a network with local spillovers. In the model, the allocator is mandated to target every unit in the network, either directly or through spillovers, and the only source of heterogeneity for units is its position in the social network. Furthermore, units from the target network are able to extract rents from the allocator. This is a common structure in political economy, in that there is one principal and many agents. The most natural applications of the model include: 1) a government trying to allocate a public good that has local spatial spillovers, such as a school, where individuals from surrounding areas may also attend the school; and 2) a leader trying to spread a message through a bureaucratic organization with incentive payments.

On the whole, the model demonstrates the existence of several realistic features that have generally not been incorporated into models of allocation. This paper shows that the setup described is only tractable when one considers the social network and, as such, suggests a revision to existing theories of allocation and rent-seeking. In the model, units are able to use their positions in the network to extract greater rents from the allocator, and some level of corruption and rent-seeking persists even in equilibrium. The model also uncovers an interesting mechanism. While units may use their social position to extract rents from the allocator, the allocator may engender competition between influential units over the social network to reduce rent-seeking and corruption.

1.2.2 Analyzing Randomized Experiments with Spillovers

This paper develops a general inferential framework for causal identification in randomized experiments in the presence of spillovers. Existing approaches focus on models of the underlying stochastic process governing spillovers or a priori knowledge of exactly which units share spillovers. This paper shows that the researcher may identify causal quantities of interest, without such strong assumptions, by analyzing the experiment with respect to inclusion probabilities induced by increasingly strong "social distance" restrictions. The social distance approach characterizes a fully general framework for causal identification in the presence of spillovers.

This paper has serious implications for practitioners of randomized experiments. First, and foremost, it describes how spillovers cause difficulties in causal inference and the minimal assumptions that need to be met for non-parametric causal identification (which is typically the goal in randomized experiments). Many political economy and policy questions of interest explicitly deal with spillovers, and sometimes the spillover itself is the quantity of interest. These questions include anything from understanding the impact of vaccination to the role of an information campaign in changing practices and attitudes. This paper shows how various causal quantities of interest may be formed and estimated efficiently, as well as providing a statistical framework that can be readily used by applied researchers.

1.2.3 A Tale of Two Villages: Kinship Networks and Preference Formation in Rural India

This study investigates the effect of kinship networks on vote choice and issue preferences over an electoral campaign in rural India. The study analyzes data collected on political preferences and kinship networks in two villages just before and after the campaign period during the 2011 Assembly election in West Bengal. The paper finds very strong kinship network effects on changes in political opinions and vote choice over a campaign.

While the previous two papers critique methods from game theory and statistical theory, this paper focuses on theoretical implications of using social networks. This paper shows that villagers engage in a far broader set of behaviors than is typically assumed in the narrative of clientelism and patronage, and these behaviors tend to be coordinated over kinship networks. Specifically, the paper finds that individuals use kinship networks to pool salient political information, engage in political discussion and coordinate political behavior. The kinship network also provides a space that is insulated from other political actors and engenders the independence of the voter. This shows that social structure and personal networks can help mitigate difficulties associated with weak state capacity, low information, lack of urbanization, and poor education in democratic developing countries and generate stable democratic practice.

1.3 Discussion

Many of the discussions in this dissertation revolve around methodological issues. For political scientists, what we say matters as much as how we say it. Standards of methodological rigor, whether in the statistical standards of causal inference or deductive standards of game theory, must generally be met; when they are not met, we must try to understand how our methods deviate from them. Unfortunately, the introduction of social networks into the analysis often complicates things. In my own work, I have rarely faced resistance to the theoretical principles of social interconnectedness affecting the things that we study. Rather, hesitation to use theories of social networks in one's analysis has stemmed from methodological difficulties. Each of these papers contains a difficult methodological problem (whether mathematical or statistical) and attempts to solve the issue in the course of the paper. The hope is that many of the methodological tools discussed in these three papers will help political scientists rigorously study a broader array of topics. The general treatment on the analysis of randomized experiments in the presence of spillovers will be of particular interest to practitioners.

At the same time, it would be a mistake to suggest that the only motivation for this dissertation is methodological. Substantively, this project aims to demonstrate how we think about politics changes radically when we move from atomistic individuals as the units of analysis to a universe where the units of study are interconnected. The introduction of social relations and social networks allow the researcher to consider a host of behaviors that cannot be easily addressed in the atomistic world. When units are connected, they have an increased capacity to cooperate and coordinate actions together. They may, in the extreme, not even choose to make decisions for themselves, choosing instead mimic the actions of a social relation. Social networks can also be used to describe hierarchy and power relations. One can characterize units that are more central in a social network, and those that are more isolated, and may even characterize the direction a spillover flows between each pair of units (as in a directed network).

While the dissertation often deals with abstract notions of social networks, my own motivation to study social networks comes from something more concrete. Much of my substantive work has focused on political behavior in India. One of the most striking features of studying the Indian system is precisely the prominent role played by larger social structures, like religion, caste, and family, in all spheres of life. The political economy literature on India has had occasion to consider, for instance, the role of caste in voting behavior and patronage, nepotism in candidate selection, and the structure of relations created as individuals migrate from rural to urban areas. Each of these major research agendas puts social relations at the center of the analysis. As I began to search for a topic to study, I found many of the existing tools in political economy inadequate to study the prominent structural concerns in Indian society, and I began to look to theories of social networks. In order to develop this perspective, many of the concepts used in the dissertation have been borrowed from sociology, which has had a long-standing engagement with the power of social networks.

However, as network theory is still relatively young in political science, I have chosen to remain agnostic on some of the central debates in sociology regarding social structure. In particular, the functionalist position (Parsons, 1951) argues that outcomes cannot be simply boiled down to individual, rational actions. Rather, the outcomes we observe are often the result of the larger social structure and the social norms that govern it. By contrast, exchange theory (Homans, 1958) argued that individual actions are impacted by the larger social structure and social networks but that all outcomes can be boiled down to decisions by individual agents. At the core of this debate is the utility of methodological individualism in explaining larger social outcomes. My paper on the impact of kinship network seems to be much more strongly in the functionalist tradition. In reality, however, the duration of the study is too short to observe changes in the kinship network, so there is room for non-functionalist interpretation of the results as the shape of kinship networks can respond to individual behavior in the long run. My game theory paper on rent-seeking over a network, by contrast, seems to be much more firmly in the tradition of exchange theory. However, in the paper, agents make fairly simple decisions and interact over a fixed social network that they cannot change, which introduces shades of functionalism.

As the literature on social networks in political science grows, we will once again begin to engage in the same theoretical debates described above. However, in order to get to that point, we must develop the tools necessary to study social networks with the demands of modern methodological rigor.

Chapter 2

On Rent Extraction and Efficient Allocation over Social Networks

Abstract

This paper develops a model of allocation and rent-seeking over a network with local spillovers. In the model, the allocator is mandated to target every vertex in the network, either directly or through spillovers, and the only source of heterogeneity for vertices is network position. Furthermore, members of the target network are able to extract rents from the allocator. The model is generally tractable, and the paper identifies several concepts to understand the level of rent-seeking and patterns of allocation over the network. In general, the allocator may use the network structure to engender competition between vertices to limit rents, while the vertices may use privileged positions in the network to extract greater rents from the allocator.

2.1 Introduction

This paper develops a model of allocation and rent-seeking over a network with local spillovers. In the model, the allocator is mandated to target every vertex in the network, either directly or through local spillovers, but the vertices in the network can extract rents from the allocator.[1]

It is shown that the allocator may use the network structure to engender competition between vertices to limit rents, while the vertices may use privileged positions in the network to extract greater rents from the allocator. The model is quite flexible and generally tractable, yielding, at times, quite surprising results about inefficiencies in allocation. This model has practical implications for those scenarios that can be modeled as "location problems," common in the study of communication networks, word-of-mouth advertising, diffusion of technology, and public goods provision. This paper is among the first to address the relationship between central allocation of goods and rent extraction in contexts with local, network-like spillovers.

In the two-period game, each vertex requests a rent, and the allocator determines the most efficient way to target the population, inclusive of rents.

Since the setting is a social network, the model simultaneously describes social heterogeneity (structure of the entire network) and variation in social status (network position for each vertex). In equilibrium, there is variation in each vertex's rent-seeking capabilities, resulting from its position in the network. However, both the allocator's optimal allocation strategy and the rents extracted in equilibrium are mediated by the entire network structure. While each vertex tries to extract as much in rents as possible, the allocator can mitigate rent extraction by inducing competition between the vertices over rents. We use this model to shed light on several key concepts:

1. The allocator never needs to allocate to vertices that are sufficiently distant from other

[1] I would like to thank Alessandra Casella, Macartan Humphreys, Matthew Jackson, Jeffrey Lax, Debasis Mishra, Massimo Morelli, Pierce O'Reilly, Arunava Sen, David Siegel, Michael Ting, Johannes Urpelainen, Timothy Van Zandt, participants at the 12th Social Choice and Welfare Conference, and seminars at Columbia University and at the Indian Statistical Institute, Delhi for useful comments and conversations in shaping this paper. All errors are my own.

(all but one) vertices. Such vertices never extract rents from the allocator. Conversely, vertices with a high status network position are able to extract rents from the allocator.

2. In many cases, the most efficient equilibrium entails the allocator paying rents.

3. More socially integrated populations (where vertices are more connected) may actually cause greater inefficiency and rent extraction in equilibrium due to reduced competition between vertices.

4. The presence of multiple high status vertices may dampen the ability of each vertex to extract rents in equilibrium due to increased competition over rents.

5. The allocator may be able to increase efficiency by reducing the number of vertices that receive allocation in order to engender competition between the vertices.

This paper proceeds in 5 sections. Section 2 describes the substantive problem and describes how social networks offer an insight into our problem, as well as stating the game in formal terms. Section 3 derives the subgame perfect Nash equilbria of the game and discusses the role of coalition-proofness in this setting. Section 5 uses examples of equilibria to shed light on the relationship, and pathologies, between allocation and rent-seeking over the network. Section 6 concludes the paper.

2.2 Framing the Problem

2.2.1 Examples

In this paper, the allocator targets a pre-defined network as efficiently as possible by either targeting vertices directly or their neighbors in the network through local spillovers. At the same time, the vertices may extract rents from the allocator. Below, a few applications of the problem considered in this paper are discussed.[2]

[2] At this point, one should be clear about the bounds of the targeting problem. A seemingly related problem, that of a profit-maximizing allocator over a network turns out to be a very different mathematical problem. Rather than attempting to target an entire population, a profit-maximizing allocator allows access to a subset of the target population up until the point where the marginal cost of providing the object exceeds

Advertising. There is a large literature on word-of-mouth advertising and viral marketing (see Goyal (2003) for a survey of the literature). Imagine that a marketer wants to spread the word about a new Facebook application over a pre-defined population, and wants to do so as cheaply as possible. The marketer faces some cost for enlisting each new person to try the application (either in money or time). One feature of Facebook is that as soon as an individual uses an application, every Facebook friend sees the application has been used on her "news feed." A network can be modeled where the vertices are individuals and a link exists between two individuals if they are Facebook friends. The goal is thus to identify the smallest set of individuals to enlist in order to make sure that each person has either used the application or heard about it on her news feed. However, if individuals realize their attractiveness to marketers, they may ask for a higher price in exchange for trying the application.

Facility Location. Facility location is a major subject in operations research (see Daskin (1995) for a much more general overview of the field), but a particular application is addressed here. Imagine a town planner wants to ensure that each person in a village (where each person is deemed to live on some small plot of land) has access to a well. The planner determines that a well will be deemed accessible by the individual if it the center of the plot with the well is within a half-kilometer of the center (i.e. centroid distance is less than or equal to a half-kilometer) of the plot of the individual. Thus, a network for this situation can be formed by defining the vertices to be plots of land, where a link exists between two plots of lands if the distance between their centers is less than or equal to a half-kilometer. The goal is thus to determine the fewest plots of land upon which to build wells in order to make sure that each inhabitant has access to a well. However, an individual with particularly well-located land may ask for greater compensation in providing the land towards the well.

Spreading Information/Propaganda. A political party wants to deliver information/rumors/propaganda among party supporters and recruits people in the population to work for the party full-time or part-time (see, for instance, Galeotti and Goyal (2009) for a

the marginal benefit in targeting more individuals. For instance, this would occur where the marginal cost of building a well exceeds the benefit to the town planner of additional people with water access.

similar logic in marketing). The party wants to make sure that each supporter receives an amount of propaganda equivalent what would come from a single person working full-time. This is a divisible allocation because the party can employ a single person full-time or, for instance, three people at one-third time, where wages are proportional to the amount of time worked. Communication between neighbors in the network is costless. The vertices are the individual party supporters and the links represent social access. The goal is make sure that each party member receives access to some minimum amount of party propaganda as cheaply as possible. However, a particularly crucial individual in the party network may ask for greater benefits in carrying out propaganda duties.

2.2.2 Relevant Literature

The question of determining an efficient allocation, as a decision-theoretic problem, is quite similar to the idea of finding key individuals in a network, which has a been a recent focus of social networks literature in economics. Galeotti and Goyal (2009) discuss a problem quite similar to this paper, where firms decide on individuals to target within the network with information about a product in order to optimize profits, under the assumption that these individuals share the information with their neighbors. However, they simplify the problem by assuming that the firms have very low information about the network, knowing only the degree distribution of vertices in the network, whereas the structure and shape of networks is central to the analysis of this paper.[3] Unlike many applications of game theory, complete information about the network is actually a significantly more mathematically demanding problem than the incomplete information case.

Ballester et al. (2006) identify a single key individual in a criminal network. Here, the authors consider a fairly general model of spillovers (substitutable and complementary) over a network and discuss how to compute each person's influence in the network using simple results from exchange theory in mathematical sociology. These measures of network influence relate to the number of paths going through the vertex. However, as this paper shows, it

[3]This allows the authors to avoid complications resulting from various structures of the network and focus on standard results from game theory (e.g., those resulting from stochastic dominance).

turns out that the problem of finding a single key individual and multiple key individuals are quite different mathematical problems.

This paper shares similarities to Bramoullé and Kranton (2007), which uses the same structure of spillovers. They model a decentralized structure where individuals invest in a locally non-excludable good where the sum of contributions of two neighbors must always equal a fixed value. This paper differs in two ways. First, this paper models the interaction of an allocator and a target population, in addition to the members of the population with each other, which allows for consideration of both the optimal allocation over a network and rent extraction from the allocator.[4] Second, without rent extraction concerns the allocator can always provide the most efficient allocation over a network. Thus, this framework allows us to consider the extent of deviations from the most efficient allocation over a network due to rent extraction. Finally, the equilibria in Bramoullé and Kranton (2007) amount to the decentralized provision of a good with local spillovers. Comparison with these equilibria allow us to make claims about the relative efficiency of decentralized provision versus centralized provision, net of rents, of goods with local spillovers.

2.2.3 Formalizing Spillovers

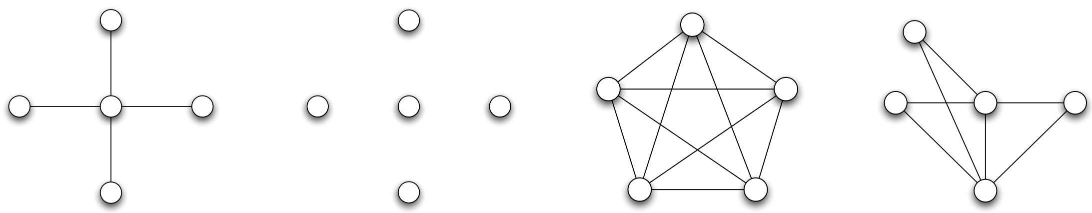
We assume that there exists some *social distance* metric over the population for a given good e, ρ_e, where the social distance between any two vertices, x and y, is represented as $\rho_e(x, y)$. Let a^* denote a vertex receiving positive allocation. A vertex, x, is deemed to be targeted if $\rho_e(a^*, x) < \overline{\rho_e}$, $\overline{\rho_e} \in \mathbb{R}$.

Let the set V, with cardinality n, denote the set of vertices in the target population, and let S denote the set of $\frac{n(n-1)}{2}$ edges that can be formed from the elements of V. We will let $E \subseteq S$ denote the set of edges over which spillovers occur, i.e. $E = \{(x, y) \mid \rho(x, y) < \overline{\rho_e},\ (x, y) \in S\}$

[4] There is virtually no literature on rent extraction in networks. The closest literature is a formal and experimental literature in sociology on exchange networks. An exchange network is a social network along which dyads of the network may bargain and make exchanges, where resources accrued to an individual in the network are seen to be a function of her network position. Within this literature there are several papers (e.g. Markovsky et al. (1988), Bienenstock and Bonacich (1993), and Bonacich (1987)) that develop axiomatic and game theoretic approaches to measures of power as a function of network position. In our model, instead of this power resulting from exchanges in the population, the power results from an increased ability to extract from the allocator, which leads to very different results.

. Using this this information, we can construct a graph (network), G, from V (vertices) and E (edges) as a representation of the structure of potential spillovers over the population.

Formally, we define the graph as a pair $G = \langle V, E \rangle$, where V will denote the vertices and E will denote the edges of the graph . We will refer to the set of vertices linked to a vertex v as the neighbors of v and denote the set as $N(v)$, where the number of neighbors for v, $|N(v)|$, will be referred to as the degree of v . Finally, it will be useful in our proofs to refer to dominating sets. A *dominating set*, $D \subseteq V$ is a set of vertices such that every vertex in V is either in D or a neighbor of a vertex in D.



(a) A Patron (b) Pure Private Good (c) Pure Public Good (d) Complex Spillovers

Figure 2.1: Some Graphs/Networks

Figure 2.1(a) represents a "star" graph with an obvious important vertex, where a good will be accessible to the whole population only if it allocated to the central vertex or to the other 4 vertices. Figure 2.1(b) represents a particularistic good with isolated vertices. Figure 2.1(c) represents a "complete" graph which represents a pure public good. Figure 2.1(d) represents a slightly more complicated graph where the good will be accessible to the entire population if it allocated to the central or bottom vertices, but not otherwise.

Two points are worth noting here. First, by definition, spillovers only move over one link in the network. If one stumbles upon a network where spillovers move over multiple links, then such a network can be rewritten in a "reduced form" where a link is drawn between any pair of vertices over which spillovers are possible. Second, this exposition uses undirected networks for clarity of exposition and intuitive appeal. However, the approach is perfectly applicable to directed network, where results regarding the degree of a vertex are replaced by the "in-degree" of a vertex, the number of vertices from which a particular vertex receives spillovers.

2.2.4 The Game

Preliminaries

The scenario is modeled as a simple two-period game, where the vertices move first, requesting rents. In the second period, the allocator determines the most efficient way to target the entire population with a local spillover good, given the requested rents. This is usually a straightforward exercise in most settings, but the problem is complicated by the fact that game takes place over network.

Let $G = \langle V, E \rangle$ be a graph where V represents vertices in the target population, and E those edges along which potential spillovers can occur. G will be assumed to be connected; that is, there exists a "path" between any two vertices in the network.[5] In the first period, the vertices, $i \in V$, request a rent or private good, $r_i \geq 0$, and the objective of each vertex will be to maximize her personal rents. The allocator will allocate a local spillover good, where the vertex that has directly received the good as well as her neighbors benefit, in a manner that targets the entire population (discussed below). Crucially, the allocator must also pay the requested rents associated with those vertices directly receiving the local spillover good (or any portion thereof). The goal of the allocator is to minimize costs of allocating this local spillover good over the entire population, inclusive of rents.[6]

Let the allocator provide a level of goods, e_i, where $i \in V(G)$. Two versions of the game will be considered, when $e_i \in \{0, 1\}$, the indivisible allocation, and when $e_i \in [0, 1]$, the divisible allocation. In the indivisible allocation, the allocator may only allocate the good as a whole or not at all, as in a building. In the divisible allocation, the allocator may allocate fractions of the good, as in scenarios where money equivalents of the good are being considered.[7] A population will be *targeted* if the sum of allocations to each vertex or a

[5] All of these results can be generalized, using the fact that all graphs are a union of connected components.

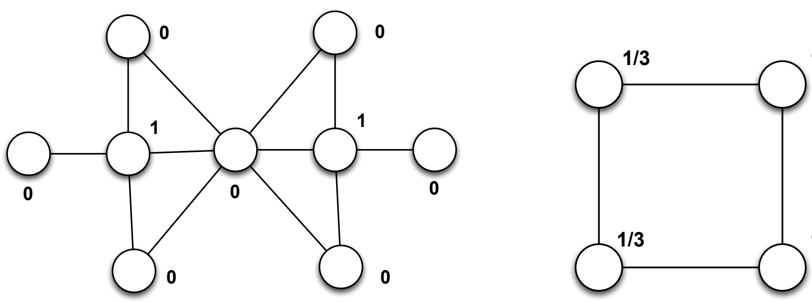
[6] In this game, we will assume that the allocator can neither receive side payments nor has ability issues; she is only interested in targeting the entire population as efficiently as possible. This allows us to focus on inefficiencies that result from the structure of the network, rather than focusing on the characteristics of the allocator.

[7] Simple examples include the central government paying for levels of pollution reduction in particular regions, where the pollution has spatial spillovers, or word-of-mouth advertising where the advertiser pays for level of effort in advertising towards the units of the population.

neighbor reaches a fixed amount (which will be normalized to 1), where the allocator aims to target the population as cheaply possible. The targeting problem for the allocator is defined as:

$$\min \sum_{i \in V} e_i \quad \text{s.t.} \sum_{j \in N(i) \cup i} e_j \geq 1 \text{ for all } i \in V \tag{2.2.1}$$

Define the space of possible allocations, $\mathbf{e}$, that satisfy the targeting constraint ($\sum_{j \in N(i) \cup i} e_j \geq 1$ for all j) as $\mathcal{E}(G)$ and the space of allocations solving the targeting problem in (2.2.1) as $\mathcal{E}^*(G) \subseteq \mathcal{E}(G)$.[8] Those vertices i receiving $e_i > 0$ will be referred to as *allocated vertices*, and as having received *positive allocation*. The elements of $\mathcal{E}^*(G)$ will be referred to as the *efficient allocation*.



(a) Bowtie

(b) Square

Figure 2.2: Some Examples of Efficient Allocations that Solve the Targeting Problem

Figure 2.2 shows examples of efficient allocations that solve the targeting problem with the numbers indicating the investment in each vertex. Figure 2.2(a) represents an *indivisible* allocation since each vertex receives either a 0 or 1. Furthermore, contrary to standard intuitions, the vertex with the highest degree does not receive investment from the allocator in the efficient allocation. Figure 2.2(b) shows a *divisible* allocation since vertices receive investments between 0 and 1. Under such a situation, we expect a more evenly distributed allocation as compared to an indivisible allocation. Efficient allocations are discussed in detail in appendix B.

[8] The setup of the targeting problem implicitly assumes anonymity, since the cost is not a function of specific vertices, and perfect substitutability, since allocating $\frac{1}{2}$ to neighboring vertices is as costly as allocating 1 to a single vertex. These assumptions allow for a focus on network intuitions in the analysis and make sure that vertices are not differentiated in any way other than by network position.

Formal Statement

Let some population distributed according to a network G, with $n = |V|$. Let e_i be the allocation to $i \in V$, with $e_i \in \{0, 1\}$ (indivisible) or $e_i \in [0, 1]$ (divisible). Let each $i \in V$ have utility $u_i(e_i, r_i)$, where $\frac{\partial u_i}{\partial r_i}, \frac{\partial u_i}{\partial e_i} > 0$.[9] The cost of the allocation to the allocator is a function of a non-negative request vector $\mathbf{r}$ and a feasible allocation $\mathbf{e}$. We define $\mathbb{C} : 2^{V(G)} \times \mathbb{R}_+^N \to \mathbb{R}_+$ as $\mathbb{C}(J, \mathbf{r}) = \min_{\mathbf{e}} \sum_J (e_i + r_i)$, where J is the set of vertices receiving positive allocation such that every vertex in the population is targeted directly or indirectly through spillovers. This represents the cost associated with efficient allocation given that a particular set of vertices receive positive allocation and a particular profile of rent requests.

Then, the allocator's problem given rents, $\mathbf{r}$, is described by selecting a set I to minimize $\mathbb{C}$:

$$\min_I \mathbb{C}(I, \mathbf{r}) \ = \ \min_I \left(\min_{\mathbf{e}} \sum_{v \in I} (e_v + r_v) \right) \quad \text{s.t.} \sum_{j \in N(i) \cup i} e_j \geq 1 \text{ for all } i \in V(G), \text{ where } I = \{v | e_v > 0\} \tag{2.2.2}$$

We will denote the space of allocations, $\mathbf{e}$, satisfying (2.2.2) as $\mathcal{E}^{\mathbb{C}}(G, \mathbf{r})$. It will be shown that the solutions to the set of allocations in the decision-theoretic problem will always be part of some equilibrium; that is, $\mathcal{E}^*(G) \subseteq \mathcal{E}^{\mathbb{C}}(G, \mathbf{r})$. It will further be shown that $\mathcal{E}^*(G)$ is identical to the set of feasible allocations satisfying the game under coalition proofness.

The game play is as follows:

1. Each $i \in V(G)$ selects requests some rent, $r_i \geq 0$.

2. The allocator selects some $\mathbf{e}$ and pays $\sum_I (e_i + r_i)$, where I is the set of vertices receiving positive allocation, subject to the constraint that

$$\sum_{j \in N(i) \cup i} e_j \geq 1 \text{ for all } i \in V(G)$$

[9] It turns out that e_i is an extraneous parameter in the vertex's utility function. Intuitively, this implies that a vertex's private valuation in the good does not affect game play.

We will assume that the allocator chooses with uniform probability between solutions that minimize cost (that is, she chooses between identical cost).

2.3 Equilibrium

In this game, the allocator will attempt to minimize the sum of the rents paid to allocated vertices and the aggregate cost of the allocation. One might believe that the vertices in the network can arbitrarily raise requests, and thus costs, to the allocator. However, this non-cooperative framework puts restrictions on this behavior, as each vertex is competing with other vertices over the network for rents.[10] Each vertex, then, is bound by the threat of the allocator choosing to allocate to other vertices, what will be called a *binding set*.

Definition 2.3.1. *For a given request vector,* $\mathbf{r}$*, and as set of allocated vertices,* I*, a set* B_i *is called a* ***binding set*** *of* $i \in I$ *if it solves*

$$B_i = \arg\min_{S} \; \mathbb{C}(S, \mathbf{r}) - \mathbb{C}(I, \mathbf{r}) \quad \textit{where } i \notin S$$

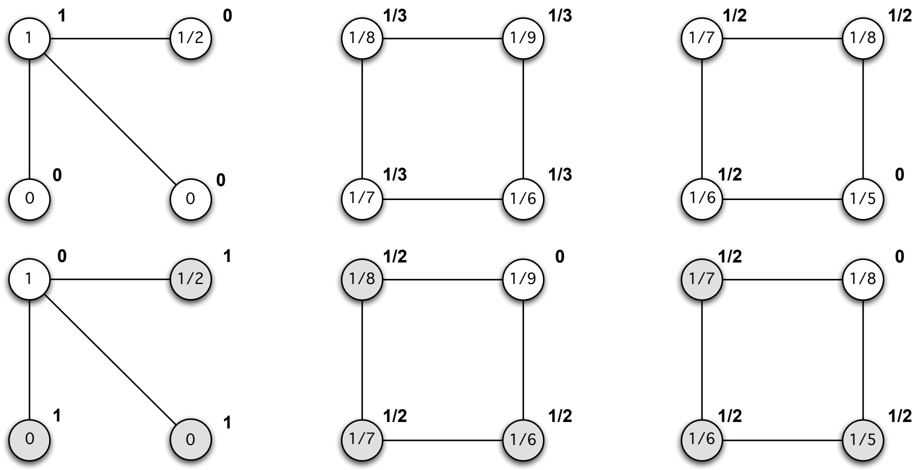
The existence of a binding set for a vertex will ensure tractability of the model since it represents a threat for the allocator to deviate by allocating to other vertices; otherwise, a vertex could extract an arbitrarily high level of rents from the allocator. Thankfully, binding sets exist quite generally, as it always exists as long as the vertex is not isolated, i.e. as long as it has neighbors in the network.[11]

Theorem 2.3.2. *If* G *has no isolated vertices, there exists a binding set for each* $i \in V(G)$*.*

Proof: The set $V(G) \setminus \{i\}$ is a dominating set in G for each $i \in V(G)$. It follows that some subset of $V(G) \setminus \{i\}$ forms a binding set for each $i \in V(G)$. $\square$

[10] Although it is not discussed in this paper, one can use the setup here to describe a transferable utility framework through potential functions. The trick is to establish an equivalence between a vertex's marginal contribution and the rent extracted in equilibrium (**??**).

[11] In the context of directed networks, the conditions for theorem 2.3.2 would be that the vertex is the target of spillovers from some other vertex in the network.



(a) Binding Set on a Star

(b) Binding Set on a Square

(c) Another Binding Set on a Square

Figure 2.3: Binding Sets

The two sets of graphs show some graphs and the binding set for a particular vertex on the graphs shaded in gray. The bold numbers outside the circles denote amounts given to allocated vertices, and the numbers inside the circles denote requests. In 2.3(a), the we show the binding set of the vertex with degree 2. In 2.3(b) and 2.3(c), we show the the binding set for the vertex in the upper right corner of the square. As the last two figures show, efficient allocation and the binding set depend upon the rent requests, even on the same graph.

In figure 2.3, we display those allocations and allocated vertices that reach the minimum cost bound for the allocator, as well as binding sets for specific vertices.[12] We see that both the allocated vertices and the binding sets are dependent upon the shape of the network and the profile of requests. The allocated vertices in figures 2.3(a) and 2.3(b) show allocated vertices and allocations that are consistent with the efficient investment over the graphs. However, if a vertex requests too much, as in 2.3(c), the allocator may be forced to select allocated vertices and allocations inconsistent with the efficient solution to the targeting problem. In figure 2.3(a) we see that if vertices with neighbors of degree 1 are allocated vertices, then those vertices of degree 1 must be included in the binding set of that vertex. Figures 2.3(b) and 2.3(c) both display binding sets for the top right vertex of the graph. Here, we see that different allocations may yield similar binding sets over the same graph.

[12]Intuitively, this a way to visualize the best response function for the allocator.

In equilibrium behavior, vertices will adjust positive requests to ensure inclusion as an allocated vertex. Thus, it should be clear that the bottom right vertex in 2.3(c) will lower her request to ensure an allocation similar to that of 2.3(b). Similarly, in 2.3(b), all vertices beside the bottom right vertex should realize that they can request more and still be included as an allocated vertex. As we will show in some detail below, equilibrium behavior entails each vertex extracting as much as possible from the allocator without giving it the incentive to shift to the allocated vertices corresponding to its binding set.

The subgame perfect Nash equilibria (SPNE) of the game can now be characterized:

Theorem 2.3.3. *Assume there exists a binding set for each $j \in J$. Then, an SPNE exists and the set of SPNE is characterized by the following conditions:*

1. *Let J be the set of vertices receiving positive allocation in equilibrium. For each $j \in J$:*

$$\mathbb{C}(J, \mathbf{r}^*) \leq \mathbb{C}(K, \mathbf{r}^*) - r_k^* \quad \textit{for all sets } K \textit{ such that } j \notin K,\ k \in K \setminus J$$

2. *Define the efficient allocation with allocated vertices J as $\mathbf{e}^*$. For each $j \in J$, define the most efficient allocation to the vertices of B_j as $\mathbf{e}^{B_j}$. The equilibrium request vector, $\mathbf{r}^*$, is any request vector satisfying the two conditions below:*

   (a) For each $j \in J$, each $k \in B_j \setminus J$ makes the request $r_k^* = 0$

   (b) For each $j \in J$:

$$\mathbb{C}(B_j, \mathbf{r}^*) = \mathbb{C}(J, \mathbf{r}^*) \Rightarrow r_j^* = \sum_{k \in B_j} e_k^{B_j} - \sum_{l \in J} e_l^* - \sum_{m \in J \setminus B_j \setminus \{j\}} r_m^*$$

Corollary 2.3.4. *If a SPNE exists, there is always a SPNE characterized by the following three conditions:*

- *The allocator solves its decision-theoretic problem to determine an allocation and the corresponding set of allocated vertices, I, i.e. $\mathbf{e}^* \in \mathcal{E}^*(G)$*

- *Each vertex $v \in V(G) \setminus I$ requests $r_v^* = 0$.*

- *Each vertex $i \in I$ requests (as in theorem 2.3.3):*

$$r_i^* = \sum_{j \in B_i} e_j^{B_i} - \sum_{k \in I} e_k^* - \sum_{l \in I \setminus B_i \setminus \{i\}} r_l^*$$

The equilibrium in corollary 2.3.4 constitutes a particularly robust equilibrium that generally exists and is not susceptible to coalitional deviations (as shown below).

2.3.1 Examples

We now discuss a few simple examples. Intuitively, a vertex connected to vertices of degree 1 will be able to extract more in rents. The examples below show the basic idea.

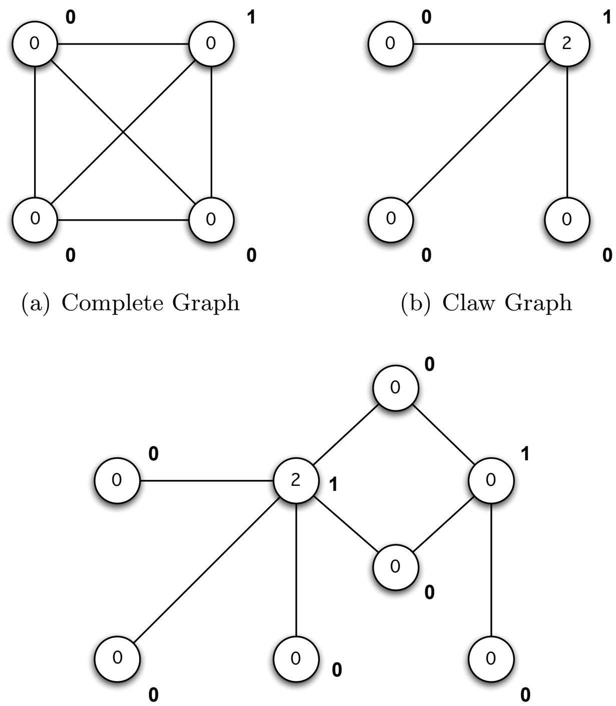
The graph in figure 2.4(c) shows that a vertex linked to other vertices of degree 1 is in a privileged position in being able to extract rents, irrespective of the rest of the graph. However, if the vertex is only linked to a single vertex of degree 1, then it might not be able to extract rents. The next, fairly straightforward, theorem makes this idea more precise.

Theorem 2.3.5. *If some vertex has two vertices of degree 1 as neighbors, it can always extract rents greater than or equal to 1 in equilibrium.*

We now discuss the SPNE of the square example that has been used in our discussion of efficient divisible allocations and binding sets above. The unique equilibrium is shown in figure 2.5.

Let us analyze equilbrium behavior of vertices in making their requests. Once again, we focus attention on the vertex on the upper right, vertex B. Sticking with our earlier convention, the efficient targeting will be over the set I and the binding set of B will be denoted as B_B.

The cost to the allocator for the graph in figure 2.6(a) is $\mathbb{C}(I, \mathbf{r}) = \frac{1}{3} \times 4 + r_A + r_B + r_C + r_D$, and cost for the graph is 2.6(b) is $\mathbb{C}(B_B, \mathbf{r}) = \frac{1}{2} \times 3 + r_A + r_C + r_D$. Thus it follows that the



(a) Complete Graph (b) Claw Graph

(c) Various Abilities to Extract Rents

Figure 2.4: Various Equilibria

Figure 2.4(a) shows an equilibrium request with all players requesting 0, since the allocator can just choose any vertex in the population whereas figure 2.4(a) shows the center of a claw graph or 3-star is able to extract positive rents. The graph in 2.4(c) shows that a vertex connected to many vertices of degree 1 is able to extract rents, even in a more complicated graph.

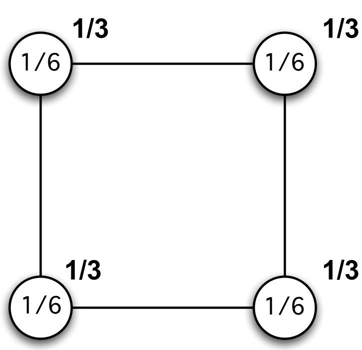
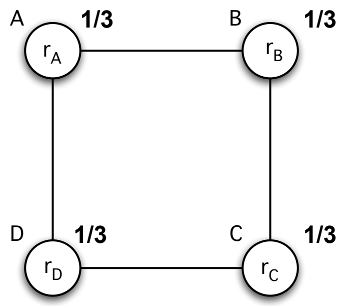
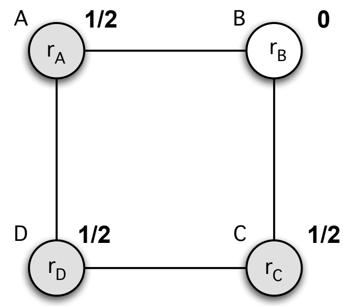


Figure 2.5: SPNE of a Square



(a) Requests on a Square



(b) Binding Set for B

Figure 2.6: Illustrating the SPNE on a Square

Figure 2.6(a) represents an arbitrary profile of requests over a square and 2.6(b) represents the binding set for B.

allocator chooses the allocation in figure 2.6(a) as long as:

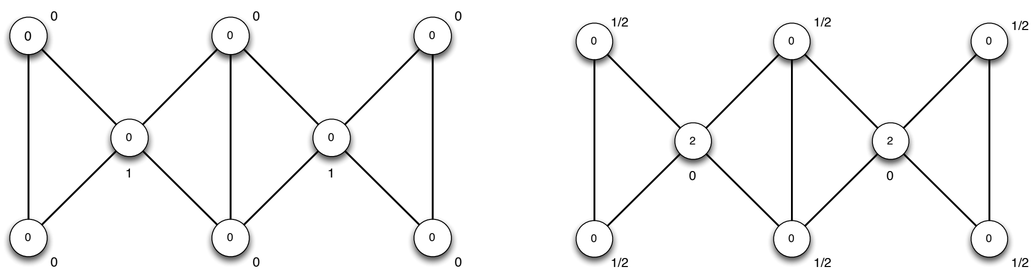
$$\mathbb{C}(I,\mathbf{r}) \leq \mathbb{C}(B_B,\mathbf{r}) \;\Rightarrow\; \frac{1}{3}\times 4 + r_A + r_B + r_C + r_D \leq \frac{1}{2}\times 3 + r_A + r_C + r_D \;\Rightarrow\; r_B \leq \frac{1}{6}$$

It follows that B is included in the allocation as long as $r_B \leq \frac{1}{6}$. Since B attempts to extract as much as possible, this inequality will bind, so $r_B = \frac{1}{6}$ in equilibrium. A similar calculation follows for each of the remaining three vertices.

The square example is interesting because it forces us to distinguish between rent-seeking behavior and social inequality. By one normative standard, the equilibrium over a square is an ideal scenario because each vertex receives an identical share of goods. On the other hand, the fact that each vertex receives positive allocation allows each vertex to extract rents from the allocator. This scenario requires us to think carefully about the tradeoff between egalitarianism and corruption. In the next section, we return to the square example to devise a scheme that will reduce rents (and increase efficiency of the equilibrium) but lead to greater inequality in goods allocated over the network.

2.3.2 Refining Equilibria

In this section, we discuss some abstract properties of the equilibria. First, we show that all equilibria may not be efficient. In this case, we define an efficient equilibrium as the lowest cost equilibrium available to the allocator in the game.



(a) Efficient Equilibrium

(b) Inefficient equilibrium

Figure 2.7: Efficient and Inefficient Equilibria

Figure 2.7(a) shows an efficient equilibrium and figure 2.7(b) shows an inefficient equilbrium over the same graph. This occurs because the two vertices receiving allocation in the efficient equilibrium must coordinate their requests together. Since Nash equilibrium logic does not allow for such coordination, we end up with an inefficient equilibrium in the right panel.

The problem in the above example is that the two vertices in the center of the diamonds (which should receive allocation in equilibrium) need to coordinate their rent requests. Since the Nash concept does not allow for this, this creates a higher cost equilibrium. We are naturally interested in how refinements might allow us to select the minimum cost equilibrium.

We might feel the solution concept of subgame perfect Nash equilibrium is not restrictive enough in that it does not allow deviations by a coalition of vertices. We investigate this refinement of the SPNE concept by allowing for some coordination while maintaining the non-cooperative environment. A refinement that makes sense here is that of a *coalition-proof Nash equilbria* (Bernheim et al., 1987) which allow subpopulations of the entire network to make self-enforcing contracts. In other words, a coalition may coordinate upon a deviating strategy if members of the coalition are made better off by the deviation, and the deviating strategy does not allow further deviation from a subcoalition. A more restrictive refinement is the concept of *strong equilibrium* (Aumann, 1959), which requires that there be no coalitional deviation which increases the utility of each vertex. In order to guarantee tractability, we will assume that that allocator prefers to allocate to the set of vertices containing the entire deviating coalition.[13] Theorem 2.3.6 shows that all coalition-proof equilibria give positive allocation to the same set of vertices and have identical costs.

Theorem 2.3.6. *The SPNE is coalition-proof if and only if* $\mathbf{e}^* \in \mathcal{E}^*(G)$*. That is, the allocated vertices in equilibrium correspond to the vertices receiving positive allocation in the efficient solution to the targeting problem if and only the equilibrium is coalition-proof. All equilibria are of identical cost to the allocator.*

Corollary 2.3.7. *Every SPNE is a strong equilibrium unless there exist some* $i, j \in I$ *under the equilibrium request,* $\mathbf{r}^*$*, where and* $i \in B_j$ *and* $j \in B_i$ *and* $\mathbb{C}(B_j, r^*) = \mathbb{C}(B_i, r^*)$*.*

2.4 Implications of Equilibrium

2.4.1 Internalizing Rent Demands

In this subsection, we investigate how the rent request of vertex is affected by the rent request of other vertices in the graph. We begin by analyzing a graph that is a hybrid of the other cases we have discussed. In this example one part of the graph constitutes a square and another part includes a vertex that is connected to two vertices of degree 1. The unique equilibrium is shown in figure 2.8.

The graphs in figure 2.9 represent the request strategy of vertex F and of surrounding vertices. We can draw two lessons about competition across vertices from this example. First, the vertices in the binding set for F that are not allocated vertices must request 0 in equilibrium. Secondly, vertices that request non-zero rents and are not in the binding set of F actually hurt F's ability to extract rents in equilibrium. The allocator will select the those vertices receiving positive allocation in the efficient solution to the targeting problem, which we denote by I.

[13]These terms are generally used in a normal form game. For our purposes, we will use these terms to refer to the induced game in the first stage from using backwards induction on the action of the allocator.

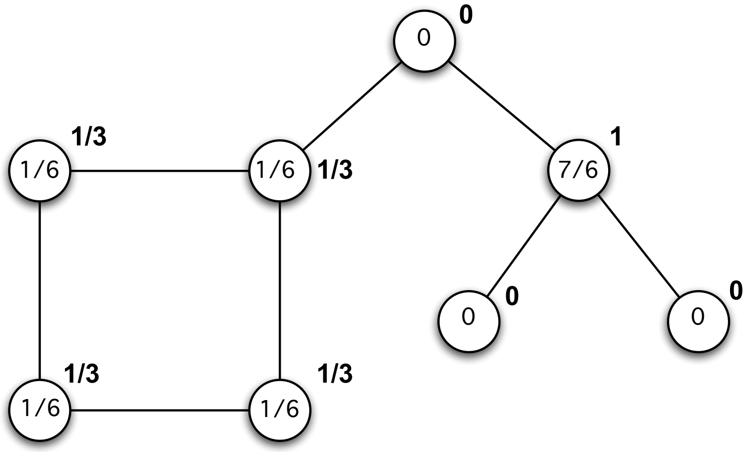
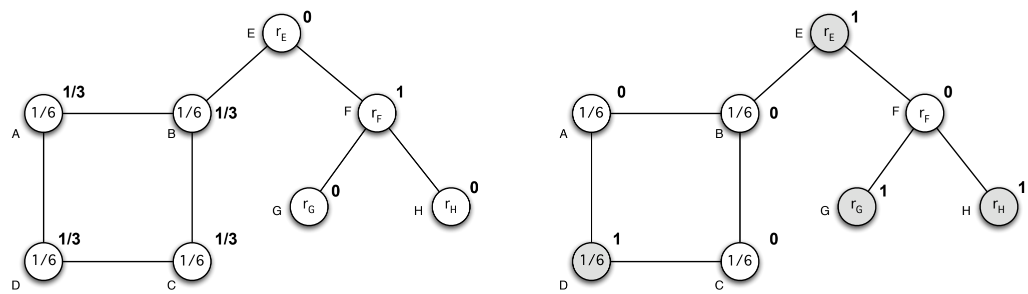


Figure 2.8: SPNE of a Graph



(a) Requests on a Graph

(b) Binding Set for F when $r_E \leq \frac{1}{6}$

Figure 2.9: Illustrating the SPNE

Figure 2.9(a) represents a profile of requests on the graph where the requests on the square are as shown in figure 2.8. Figure 2.9(b) represents the binding set for vertex F.

In figure 2.9(a) and 2.9(b) , we show a profile of requests where the requests of E, F, G, and H are arbitrary except for the fact that $r_E \leq \frac{1}{6}$ (so that we have the correct binding set).[14] The cost to allocator for allocation to the allocated vertices is $\mathbb{C}(I, \mathbf{r}) = \frac{1}{3} \times 4 + \frac{1}{6} \times 4 + 1 + r_F$. The cost to the allocator for allocation to the binding set of F is $\mathbb{C}(B_F, \mathbf{r}) = 1 \times 4 + \frac{1}{6} + r_E +$

[14]We will also show that $r_E = 0$ in equilibrium, so the assumption does not actually rule out any equilibria.

$r_G + r_H$. Then, the allocator selects I if:

$$\mathbb{C}(I, \mathbf{r}) \leq \mathbb{C}(B_F, \mathbf{r}) \;\Rightarrow\; r_F \leq \frac{7}{6} + r_E + r_G + r_H$$

Once again, in equilibrium, the constraint will bind, so $r_F = \frac{7}{6} + r_E + r_G + r_H$. We now show that $r_E = r_G = r_H = 0$ in equilibrium. Assume, towards a contradiction and without loss of generality, that $r_E > 0$. There exists some r'_E such that $0 < r'_E < r_E$. Then we have $r_F = \frac{7}{6} + r_E + r_G + r_H$, but $r_F > \frac{7}{6} + r'_E + r_G + r_H$. Hence, E has an incentive to deviate and request r'_E since this would give her some rents where as the allocation to I would give her no rents. This shows that vertices in a binding set but not in the set of allocated vertices will request no rents in equilibrium. Thus, $r_F = \frac{7}{6}$.

The difference in the costs of the allocations (i.e. the sum of the e_i terms) to I and B_F is $1 \times 4 - \left[\frac{1}{3} \times 4 + 1\right] = \frac{5}{3}$. However, we see that F is only able to extract $\frac{7}{6}$ in equilibrium. The reason this occurs is because vertices A, B, $C$and D all request $\frac{1}{6}$. Thus, while cost of allocating to the binding set is significantly higher, the allocator saves by not having to pay the total $\frac{2}{3}$ in rents resulting from the requests of A, B, C and D. This shows that the extraction power of an allocated vertex is hurt by not having other allocated vertices that can extract positive rents in its binding set.

Without network structure, this game operates under two conflicting logics. On one hand, a group that knows the allocator cannot credibly exit the negotiations can extract large rents from the allocator by behaving cooperatively. On the other hand, if the allocator knows that it will only have to allocate to some subset of the population (in order to target everyone), then it can induce competition between vertices to make sure there is no rent extraction. Network structure allows us to reconcile these logics to get a tractable non-cooperative solution in equilibrium. In particular, allocators are able to use competition between vertices to keep rent extraction down. At the same time, vertices are able to take advantage of their position in the network structure in order to extract rents from the allocator because she cannot exit.

A vertex that receives allocation in equilibrium experiences two factors that curb her ability to extract rents in equilibrium. First, it faces competition from vertices that do not receive allocation in equilibrium; such vertices lower their rents as much as possible to be attractive to the allocator (but ultimately still fail to receive allocation). The second factor is more interesting. Other vertices that receive allocation in equilibrium, but are not a part of the binding set, also limit rents. This implies that beyond competition between vertices, the mere presence of other powerful vertices in the population can limit rent extraction.

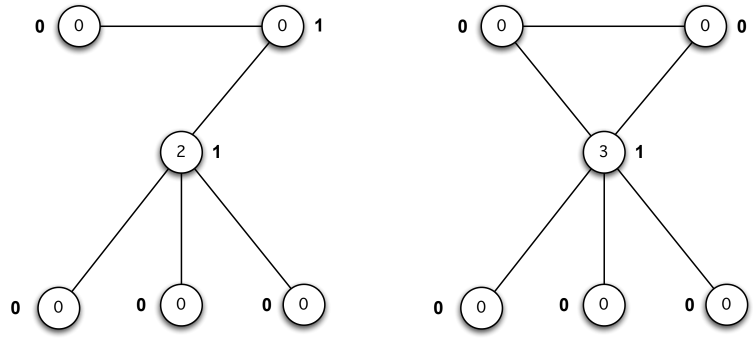
2.4.2 Adding Links

The next two subsections deal with the notion of equilibrium efficiency. First, we look at the situation where we add links to the graph. We might believe that as we add links to the graph, we should move "monotonically" towards a more efficient solution, since a complete graph represents a globally non-excludable good (and the most efficient allocation possible). However, as we show, adding links to graph may increase the ability of a vertex to extract rents in equilibrium, and we may actually create a more inefficient equilibrium if adding links induces less competition between vertices over rents.

It is easy to see that adding links never decreases the efficiency of the investment for the allocator. On the other hand, we might think that adding links decreases rents in equilibrium. This turns out to be false. The simple example in figure shows that rents might actually *increase* in equilibrium after adding links.

Simply adding links to a graph may actually strengthen a vertex vis-a-vis other vertices in the graph because adding a link to a vertex always increases its rent extraction position. Thus, adding links can induce countervailing effects, increasing the efficiency of the targeting but also increasing rent extraction. As the next example shows, adding links to a graph may actually induce *less efficiency* in equilibrium.

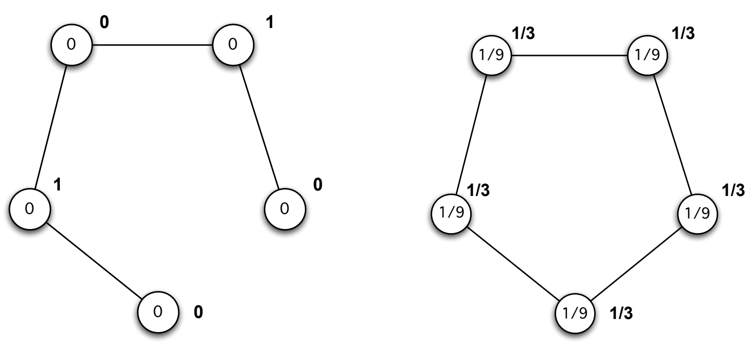
This subsection shows that while adding links to graph may create greater efficiency in allocation, there may be countervailing effects in terms of rent extraction which could potentially lead to more inefficient outcomes. First, links can strengthen the rent extraction capabilities of a vertex relative to other vertices. Second, adding links may actually reduce competition for rents between vertices, which is necessary to exert a downward pressure on



(a) Initial Graph (b) Adding a Link

Figure 2.10: Increasing Rents by Adding a Link

Figure 2.10(b) adds a link to figure 2.10(a), but the rents increase in equilibrium. Adding links may actually increase allocative imbalances.



(a) Initial Graph (b) Adding a Link

Figure 2.11: Decreasing Efficiency by Adding a Link

Figure 2.11(b) adds a link to figure2.11(a), but the increase in rents more than offsets in the increase in efficiency of the allocation. In this case, adding a link decreases the competition between vertices over rents.

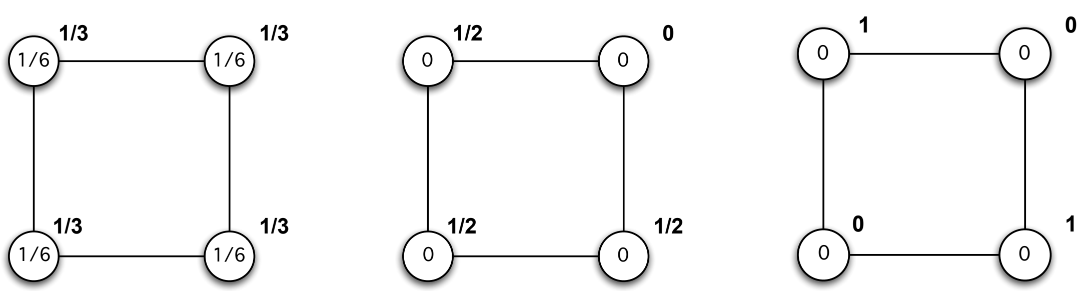
rent extraction.

2.4.3 Domination Efficiency

In this subsection, we discuss a mechanism to potentially increase efficiency and decrease rent extraction. In particular, we introduce a useful concept, *domination efficiency*,[15] to determine when the allocator should restrict the number vertices receiving positive allocation. The intuition here is that the allocator may induce competition between vertices for rents by limiting the number of vertices receiving positive allocation.

Definition 2.4.1. *A **k-game** is a game where the allocator is restricted to give positive allocation to at most k vertices in the graph G. In other words, the allocator selects the most efficient allocation when positive allocation is restricted to at most k vertices in the graph, subject to the restriction that each vertex in the population meets the targeting constraint.*

In order to see why the allocator might want to restrict the number of vertices, consider the example of the square. By potentially allowing for positive allocation to all 4 vertices, i.e. a 4-game, the allocator actually ends up spending more in equilibrium. In the 3-game, the allocator gives an allocation of $\frac{1}{2}$ to each of 3 vertices. But since only 3 vertices will receive allocation, the vertices compete to included as allocated vertices. In the 2-game, the allocator once again ends up spending as much as in the 4-game.



(a) The Equilibrium in a 4-game

(b) An Equilibrium in a 3-game

(c) An Equilibrium in a 2-game

Figure 2.12: Illustrating k-games on a Square

[15]We refer to this type of efficiency as domination efficiency because all feasible allocation give positive allocations to a dominating set.

Looking at the examples in figure 2.12, we see that the the 4-game shown in figure 2.12(a) and the 2-game shown in 2.12(c) yield an aggregate cost of 2 to the allocator, whereas the 3-game depicted in figure 2.12(b) yields an aggregate cost of $\frac{3}{2}$ to the allocator. By simply guaranteeing that one vertex will not receive allocation in equilibrium, the allocator is able to engender competition between the vertices because one vertex will be left out. Such a vertex is in the binding set of all of the other vertices, and so must request 0, according to earlier arguments. Since the argument is the same for each of the 4 vertices, each vertex must request 0 in equilibrium. The following definition relates to the cost to the allocator of various k-games.

Definition 2.4.2. *A k-game is* ***domination efficient*** *if it contains the minimum cost equilibrium for the allocator over all equilibria in all k-games.*

In the square example, the 3-game is domination efficient. The natural question is when should the allocator be restricted in the number of vertices receiving positive allocation. In other words, when does a graph with n vertices have a domination efficient n-game? Here, we provide some simple sufficient conditions.

Theorem 2.4.3. *Let all $\mathbf{e}^* \in \mathcal{E}^*(G)$ be such that $e_i^* > 0$ for all $i \in V(G)$ (in the n-game with $n > 1$), and let each i have a binding set with $n-1$ vertices. Then, the n-game is not domination efficient.*

The condition is here is quite straightforward. If each vertex receives positive allocation in equilibrium in the n-game and is bound by the rest of the vertices, then the equilibrium is not domination efficient. The basic intuition is that in such a situation each vertex requests lower rents when the number of vertices receiving allocation is lowered to $n-1$, so this leads to inefficiency in the n-game.[16]

It is trivial to show that in the indivisible game the n-game is always domination efficient because the solution entails allocating to as few vertices as possible. The implications for

[16] We could have used a slightly more general condition that states that the binding set of any allocated vertex is a subset of the allocated vertices, and allocation to the binding set is strictly less efficient than the efficient allocation. The proof would be a bit more complicated.

these results for rent-seeking and efficiency are quite interesting. Normatively, we might find it appealing that all vertices receive positive allocation (and perhaps an identical amount). However, since all vertices receive positive allocation, they can extract larger rents from the allocator. These results suggest that the allocator can decrease rents and increase efficiency by limiting the number of allocated vertices allowed, thereby manufacturing competition between the vertices for rents and allocation.

2.5 Conclusion

This paper models the interplay between rent extraction and efficient allocation over networks. Many political economy models talk of a tradeoff between corruption and efficient provision, but explicitly modeling the spillover structure shows that certain levels of rents may be necessary for more efficient provision. The model is used to understand several important principles in equilibrium behavior:

- **Heterogeneous Vertices** Key vertices in the network, in terms receiving positive allocation and being able to extract rents, are those vertices that are linked to vertices of degree 1. In particular, such vertices are able to leverage their unique ability to access a remote vertex to personal benefit. This leads to the counterintuitive result that vertices with high degree may, at times, be particularly weak in terms of targeting and rent extraction because a lot of links means that these vertices can be accessed more easily through spillovers from other vertices.

- **Adding Links.** Adding links to a graph always yields a more efficient targeting, but this can be mitigated by higher rents. In fact, one may actually induce less efficient outcomes with much higher rents by adding links. Adding links to vertices of degree 1 may cause the allocator to spread its allocation around the network, leading to less inequality in public goods allocation over the network.

- **Competition.** There is a network relationship over rents because vertices must internalize the other requested rents. In particular, if there exist some vertices that do not receive positive allocation in equilibrium, they request a rent of 0. These zero requests dampen the ability of vertices receiving allocation to request higher rents. This also guarantees that the allocator only needs to worry about the most efficient public goods allocation as a decision-theoretic problem, and the rents will sort themselves out.

- **Internalizing Rents.** One of the peculiar results in this model is that multiple influential vertices can dampen rent extraction, even though such vertices may not be direct competition with each other. In particular, the ability of an allocated vertex to extract rents is dampened by the rent extraction of other allocated vertices that are not a part of its binding set.

- **Corruption and Rents.** In this setup, in many cases the most efficient outcome will entail the allocator paying rents. While rents are dampened by competitive effects, some vertices are able to extract rents in equilibrium due to heterogeneity resulting from network position. However, at times, the allocator can induce competition between vertices by limiting the number of allocated vertices. This competition may lead to fewer rents being paid and greater overall efficiency. Scenarios where every vertex receives positive allocation will invariably lead to higher rent extraction. Thus, there is a tradeoff between egalitarian allocation of the good and levels of rent extraction.

- **Tractability.** The setup discussed in this paper is tractable as long as there are no isolated vertices (and as long as every vertex is the target of some spillover in the directed case). One might analogize our game to an auction with multiple buyers. However, in a classical auction setup, if individuals are not differentiated and not everyone receives a good, they will never extract a rent. On the other hand, if everyone receives a good in the auction, then it is not tractable. Because binding sets exist quite generally, one is able to adjudicate between these countervailing pressures, as well as heterogeneity in network position, to find non-zero rents in equilibrium even when each vertex receives positive allocation.

There are several directions for further research on efficient allocation over a network with spillovers; three are highlighted here. First, one can consider significantly more complicated interactions in a repeated setup where network changes endogenously as a function of the allocation. Here, both vertices in the network and the allocator would have to account for future network effects of current allocations and requests for rents. A second direction would be to have the allocator maximize a utility function that does not depend on complete targeting of the network, where the utility depends on which or how many vertices have been targeted and the extent of spillovers. Finally, one may construct a richer game that incorporates a decentralized allocation structure (e.g., Bramoullé and Kranton (2007)) with the centralized allocation structure in this paper to provide a nuanced analysis of provision of goods with local spillovers.

Appendix A: Mathematical Appendix

Proof of Theorem 2.3.2:

The set $V(G) \setminus \{i\}$ is a dominating set in G for each $i \in V(G)$. It follows that some subset of $V(G) \setminus \{i\}$ forms a binding set for each $i \in V(G)$. □

Proof of Theorem 2.3.3:

Let J be the set of allocated vertices in equilibrium. The proof requires reasoning about when the allocator can and cannot defect (by moving to a cheaper allocation, inclusive of rents) due to a defection by one of the vertices. Consider the first condition:

$$\mathbb{C}(J, \mathbf{r}^*) \leq \mathbb{C}(K, \mathbf{r}^*) - r_k^* \quad \text{for all sets } K \text{ such that } j \notin K,\ k \in K \setminus J \tag{2.5.1}$$

This is a direct statement of the "single defection rule" for non-allocated vertices, i.e., no single vertex that does not receive allocation in equilibrium can defect by requesting a lower rent in way that induces the allocator change her allocation by flipping the inequality in 2.5.1.

Assume the existence of a binding set. The cost inequality must bind for the binding set. To see this, consider some vertex j and request profile $\overline{\mathbf{r}}$ such that:

$$\mathbb{C}(J, \overline{\mathbf{r}}) < \mathbb{C}(B_j, \overline{\mathbf{r}}) \tag{2.5.2}$$

Choose some ε such that $\mathbb{C}(B_j, \overline{\mathbf{r}}) - \mathbb{C}(J, \overline{\mathbf{r}}) > \varepsilon > 0$. Then, the vertices $i \in J$ may defect by requesting $r_i' > r_i^*$ such that $\sum_{j \in J}(r_i' - \overline{r}_i) < \varepsilon$ That is, the allocated vertices extract more without changing the allocation decision.

Then, there exists a defection $r_j^* > \overline{r}_j$ which preserves the inequality, so j can extract more. If the constraint binds, so $\mathbb{C}(J, \overline{\mathbf{r}}) = \mathbb{C}(B_j, \overline{\mathbf{r}})$, any defection ($r_j^* > \overline{r}_j$) would imply that $\mathbb{C}(J, \mathbf{r}^*) > \mathbb{C}(B_j, \mathbf{r}^*)$ and cause j to receive no allocation (and thus no rents).

This shows that in equilibrium, $\mathbb{C}(J, \mathbf{r}^*) = \mathbb{C}(B_j, \mathbf{r}^*)$ for $j \in J$. Now assume, that there exists some vertex in the binding set but not receiving positive allocation, i.e. in $k \in B_j \setminus J$

with equilibrium requested rent $r_k^* > 0$. Since it must be the case that $\mathbb{C}(J, \mathbf{r}^*) = \mathbb{C}(B_j, \mathbf{r}^*)$ in equilibrium, there exists $0 < \overline{r}_k < r_k^*$ that induces $\mathbb{C}(J, \overline{\mathbf{r}}) > \mathbb{C}(B_j, \overline{\mathbf{r}})$, which forces the allocator to allocate to B_j instead of I. This shows that in equilibrium $r_k^* = 0$ for all $k \in B_j \setminus J$, the first sub-condition of condition 2. The second sub-condition follows from putting the result of the first sub-condition into the equality constraint shown above. $\square$

Proof of Corollary 2.3.4:

We know that for each $i \in I$, $\sum_{j \in I} e_j \leq \sum_{k \in K} e_k$ for all sets such that $i \notin K$. Our conditions imply $\mathbb{C}(I, \mathbf{r}^*) \leq \mathbb{C}(K, \mathbf{r}^*)$, which is consistent with condition 1 in theorem 2.3.3, so the allocator has no incentive to defect. The requested rents imply that the constraint binds for some K, namely B_i, so no vertex has the incentive to defect either. $\square$

Proof of Theorem 2.3.5:

Consider $v \in V(G)$ and $u, w \in N(v)$, where $N(u) = N(w) = \{v\}$. Let I be the set corresponding to an efficient allocation. Notice that I must contain v since any set of feasible allocated vertices without v must contain u and w, each allocated 1. Let J be a dominating set containing u and w. Clearly $J \setminus \{u\} \setminus \{w\} \cup \{v\}$ (with v allocated 1) is a also a feasible allocation and strictly more efficient, showing that each efficient allocation contains v. Now consider some request, R, and let I' be some dominating set not containing v (and thus containing u and w). Notice that for $r_v \leq r_u + r_w + 1$, we have $\mathbb{C}_R^I \leq \mathbb{C}_R^{I'}$. Thus, in equilibrium, $r_v \geq 1$. $\square$

Proof of Theorem 2.3.6 and Corollary 2.3.7:

In order to prove the theorem and corollary, we proceed in three steps. First, we show that any set of allocated vertices, J, that does not correspond to the efficient allocation is susceptible to deviations from some coalition. Second, we will show that the set of allocated vertices corresponding to the efficient allocation does not admit a self-enforcing deviation from a coalition (which will also show the conditions under deviations by coalitions, which are not self-enforcing, are possible). Finally, we will show that all coalition-proof equilibria admit the same aggregate cost to the allocator.

First, we show that any equilibrium with positive allocation to J such that the corresponding investment vector $\mathbf{e}^J \notin \mathcal{E}^*(G)$ is susceptible to coalitional deviations. Let us also denote the set I as the set of allocated vertices corresponding to some vector under the efficient solution to the targeting problem, $\mathbf{e}^* \in \mathcal{E}^*(G)$.

Notice that for any set J, and request vector, $\mathbf{r}$:

$$\mathbb{C}(J, \mathbf{r}) = \sum_{j \in J} e_j^B + \sum_{I \cap J} r_j + \sum_{J \setminus I} r_j$$

In order for $\mathbb{C}(J, \mathbf{r})$ to be more efficient, we need:

$$\sum_{j \in J} e_j^B + \sum_{I \cap J} r_j + \sum_{J \setminus I} r_j \leq \sum_{i \in I} e_i^* + \sum_{I \cap J} r_j + \sum_{I \setminus J} r_j$$

Since $\sum_{j \in J} e_j^B > \sum_{i \in I} e_i^*$, there exists some number $\varepsilon > 0$ such that $r_k = \varepsilon$ for all $k \in I \setminus J$ makes the right and left sides of the expression equal. By assumption, this induces the allocator to allocate to vertices in I. Thus, any allocation to J is susceptible to deviations from some coalition.

Now, we show that an allocation to the set of vertices in the efficient allocation is coalition-proof. In order to find coalition-proof equlibria, we show that we need only need to consider defections that contain some $i \in I$ and some member of the binding set $j \in B_i$. Remember that the space of subgame perfect Nash equilibria is characterized by the conditions in theorem 2.3.3. Thus, if $i \in I$ increases r_i^* without a defection from some member in B_i, $\mathbb{C}(B_i, \mathbf{r}^*) < \mathbb{C}(I, \mathbf{r}^*)$, ruling out such defections. We now consider two cases, defections which include (1) $i \in I$ and $j \in B_i \setminus I$, and (2) $i, j \in I$, $i \in B_j$ and $j \in B_i$, so an equilibrium allocation to J is susceptible to deviations from a coalition.

We now show I is not susceptible to self-enforcing coalitional deviations.

In the first case, let us consider a defection by a coalition S that contains $i \in I$ and some member of the binding set not in I, $j \in B_i \setminus I$. Any potential defection induces $\mathbb{C}(B_i, \mathbf{r}^*) \geq \mathbb{C}(I, \mathbf{r}^*)$, in which case the deviation does not increase the utility for j, or $\mathbb{C}(B_i, \mathbf{r}^*) < \mathbb{C}(I, \mathbf{r}^*)$, in which case the deviation does not increase the utility for i.

In the second case let us consider a defection by a coalition S that contains $i, j \in I$ where $i \in B_j$ and $j \in B_i$. In other words, each allocated vertex is in the other vertex's binding set. As before, it never makes sense for either i or j to decrease its requested rent. Let i and j defect to $r_i' > r_i^*$ and $r_j' > r_j^*$ respectively yielding a request vector, $\mathbf{R}'$. This implies that $\mathbb{C}(B_i, \mathbf{r}^*) < \mathbb{C}(I, \mathbf{R}')$ and $\mathbb{C}(B_j, \mathbf{R}') < \mathbb{C}(I, \mathbf{R}')$. If $\mathbb{C}(B_j, \mathbf{R}') < \mathbb{C}(B_i, \mathbf{R}')$ or $\mathbb{C}(B_i, \mathbf{R}') < \mathbb{C}(B_j, \mathbf{R}')$ then the deviation does not benefit i or j, respectively.

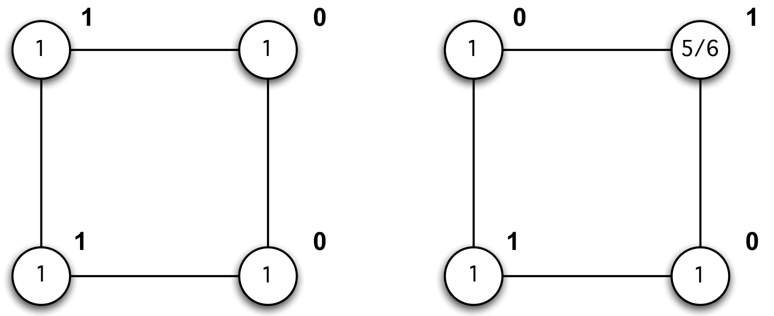
But, what if $\mathbb{C}(B_i, \mathbf{R}') = \mathbb{C}(B_j, \mathbf{R}')$? Since $i \notin B_j$ and $j \notin B_i$, this implies that i and j will be allocated vertices with probabilities p_i and p_j, which are less than 1 (because the allocator chooses randomly between options). Then, as long as $p_i r_i' > r_i^*$ and $p_i r_i' > r_i^*$, such a deviation improves the utility of both i and j. But, now notice that there exists some r_i'' such that $p_i r_i' < r_i'' < r_i'$ which induces $\mathbb{C}(B_i, \mathbf{R}') < \mathbb{C}(B_j, \mathbf{R}')$. Thus, such a coalition defection is subject to single defection by either i or j.

Finally, we show that all coalition-proof equilibria yield the same aggregate cost to the allocator, net of rents. Let I and I' denote two sets of allocated vertices that are selected in a coalition-proof equilibrium under the rent vectors $\mathbf{r}^*$ and allocations $\mathbf{e}^*$ and $\mathbf{e}'$, respectively. The previous argument shows that the targeting to I and I' yield the same, efficient outcome; thus, $\sum_{i \in I} e_i^* = \sum_{j \in I'} e_j'$. Now assume towards a contradiction that $\mathbb{C}(I, \mathbf{r}^*) < \mathbb{C}(I', \mathbf{r}')$. Then,

$$\sum_{i \in I} e_i^* + \sum_{j \in I \cap I'} r_j^* + \sum_{k \in I \setminus I'} r_k^* < \sum_{i \in I'} e_i' + \sum_{j \in I \cap I'} r_j^* + \sum_{k \in I' \setminus I} r_k^* \tag{2.5.3}$$

Using the fact that the targeting to I and I' is identical, we have $\sum_{k \in I \setminus I'} r_k^* < \sum_{k \in I' \setminus I} r_k^*$. If $\sum_{k \in I \setminus I'} r_k^* > 0$, then there exist $\varepsilon > 0$ such that $\sum_{k \in I' \setminus I} \varepsilon = \sum_{k \in I \setminus I'} r_k^*$. This would imply that the allocator would allocate to I' making I susceptible to deviations from some coalition, contradicting our previous argument. Thus, it must be that $\sum_{k \in I \setminus I'} r_k^* = \sum_{k \in I' \setminus I} r_k^* = 0$. Therefore, $\mathbb{C}(I, \mathbf{r}^*) = \mathbb{C}(I', \mathbf{r}^*)$. Now consider some other equilibrium rent vector, $\mathbf{r}'$. Direct calculation from theorem 2.3.3 shows that $\sum_{j \in I \cap I'} r_j^* = \sum_{j \in I \cap I'} r_j'$. Appealing to equation 2.5.3 and the previous arguments show that all coalition-proof equilibria yield the same cost. $\square$

Proof of Theorem 2.4.3:



(a) Potential Payoff with Coalitional Defection

(b) B Defection to Guarantee Allocation

Figure 2.13: Non-Existence of a Strong Equilibrium

Figure 2.13(a) represents a Pareto-improving defection from equilibrium for all four vertices, where the allocator randomly selects two vertices and allocates 1 to each. Figure 2.13(b) shows that vertex B can further defect to guarantee selection as an allocated vertex.

Let us consider an n-game. The condition that all vertices receive positive allocation in the efficient allocations guarantees that the cost of any binding set for a vertex $i \in V(G)$, B_i with optimal allocation $\mathbf{e}^{B_i}$, is less efficient than the most efficient allocation $\mathbf{e}^*$, so $\mathbb{C}(B_i, \mathbf{r}^*) > \mathbb{C}(V(G), \mathbf{r}^*)$. It follows that

$$r_i^* = \mathbb{C}(V(G), \mathbf{r}^*) - \mathbb{C}(B_i, \mathbf{r}^*) = \sum_{j \in V(G)} e_j^{B_i} - e_j^* > 0$$

We can then rewrite $\mathbb{C}(V(G), \mathbf{r}^*)$ as:

$$\mathbb{C}(V(G), \mathbf{r}^*) = \sum_{j \in V(G)} e_j^{B_i} + \sum_{k \neq i} r_k^*$$

Now consider the $(n-1)$-game, and let the government select the most efficient allocation to a set J with equilibrium requests $\mathbf{r}'$. If this allocation excludes positive allocation to some i', then this allocation goes to the binding set of i', so $J = B_{i'}$. Notice that the binding set for each $j \in V(G)$ is still feasible for the government. Now, for each $j \neq i'$, we have, following above, that:

$$r'_j = \mathbb{C}(B_{i'}, \mathbf{r}') - \mathbb{C}(B_j, \mathbf{r}') = \sum_{j \in V(G)} e_j^{B_j} - e_j^{B_{i'}} > 0$$

But, since $\sum_j e_j^{B_{i'}} > \sum_j e_j^*$, we know that $r'_j < r^*_j$ for $j \neq i'$. It follows that:

$$\mathbb{C}(B_{i'}, \mathbf{r}') = \sum_{j \in V(G)} e_j^{B_{i'}} + \sum_{k \neq i'} r'_k < \sum_{j \in V(G)} e_j^{B_{i'}} + \sum_{k \neq i'} r^*_k = \mathbb{C}(V(G), \mathbf{r}^*)$$

Thus, there is a more efficient outcome in the $(n-1)$-game as compared to the n-game.

□

Appendix B: Results about the Efficient Allocation

Indivisible Allocation

In this section, the paper investigates various aspects of the efficient allocation in the allocator's targeting problem, essentially the second stage of the game. The results in this section are akin to writing down the objective function of the allocator and solving for the allocation that maximizes her utility. But, in most models, this is fairly easy to do using standard methods in arithmetic and calculus. Here, however, since our analysis is over a network, these methods are of limited value, but the complicated structure allows us to develop a deeper intuition for how allocation is affected by social structure.

In order to develop intuition, we begin with the special case where $e_i \in \{0, 1\}$, what we call an indivisible allocation; that is, we investigate the some properties of $\mathcal{E}^*(G)$ when the allocator can only make a binary allocation. In section B.2, we consider how the efficient allocation changes if we allow for divisible allocations, i.e. $e_i \in [0, 1]$. We will show that there will always be a solution to the targeting problem that excludes positive allocation to vertices of degree 1 and a set of necessary conditions for the efficiency of divisible allocations.

We begin our discussion by interpreting the use of spillover networks in allocation problems. Consider an allocation in $\mathcal{E}^*(G)$ that satisfies the (indivisible) targeting problem. Form a subgraph by connecting each vertex to all possible sources with a link, i.e. for each vertex not receiving positive allocation, connect the vertex to all neighbors with positive allocation with a link. This will lead to a graph (not necessarily connected) that "spans" the network. This is a nice way to visualize feasible solutions to the indivisible allocation problem, and it generally known as a *star cover* of the network.

Definition 2.5.1. *The spillover network associated with a graph G and induced by an allocation A, $G[A]$, is the subgraph formed from the set of dyads including a vertex with positive allocation. Formally, $G[A] = (V, S)$ where $S \subseteq E(G)$ and $S = \{(u, v)|\ e_u + e_v \geq 1\}$ under A.*

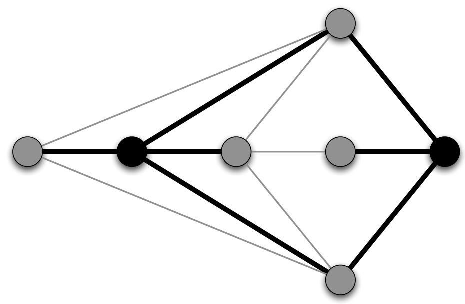


Figure 2.14: Spillover Network

Figure 2.14 shows a network with an overlaid spillover network induced by an allocation, represented by the thick black links. The black vertices represent those vertices receiving an allocation of 1, with the gray vertices receiving 0.

The question of a minimum cost allocation to the network, then, is a question of covering the network with the fewest stars, or the minimum star cover. This problem is identical to finding a minimum dominating set over the network, a classic problem in mathematical graph theory (see Haynes et al. (1998) for details). Although, intuitively, one might believe the efficient strategy is to target the vertex with the most neighbors, this intuition turns out be false, as shown in figure 2.15, where the vertex in the center has the most neighbors. The problem is that a vertex with many neighbors can reach many others but can be reached by many others as well. The example in figure 2.15 also shows an important property of the efficient allocation, discussed in section B.2, that there is always an efficient allocation that allocates nothing to vertices of degree 1 and 1 to their neighbors.

Allocation strategies can be similar on very different graphs. We tend to think of complete networks (where every vertex is connected to every other vertex) and stars (where everyone is linked to one vertex and no one else) as very different social structures, representing systems with no hierarchy and severe hierarchy, respectively. But, from the standpoint of allocation, they yield very similar results.

The difference between the two graphs in figure 2.16 does not yield a difference in allocation strategy; as shown in the paper, these graphs lead to very different rent extraction strategies.

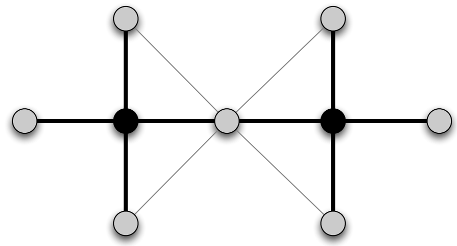
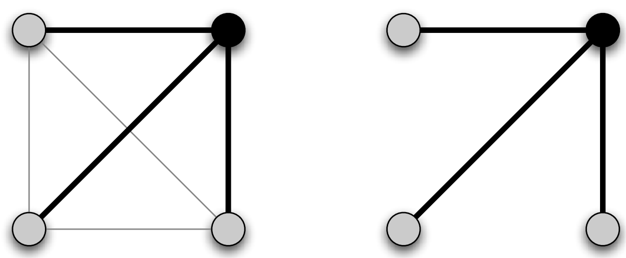


Figure 2.15: Finding the Efficient Allocation

The graphs above depict an optimal allocation. In figure 2.15, we see an example where the optimal allocation does not include giving a "1" to the vertex with the most neighbors, the vertex in the center of the graph. We have overlaid the flow network, denoted by the bold links.



(a) Complete Network (b) Claw

Figure 2.16: Complete Network and a Claw

Figure 2.16(a) shows a complete network with an efficient spillover network (bold links) overlaid on the graph, and figure 2.16(b) shows a claw (star) with the same spillover network.

Allocation Strategies

We now discuss the case where the allocation is divisible, i.e. $e_i \in [0, 1]$.

First, we show that it is always efficient to not allocate anything to vertices of degree 1, and to allocate the maximum amount (i.e. allocate 1) to the neighbors of vertices of degree 1.[17] This implies that the vertices that are connected to hard to reach areas are those vertices that can expect to receive the highest levels of allocation.

Theorem 2.5.2. *If u is a vertex of degree 1 with neighbor v, so $N(u) = \{v\}$, then there*

[17]The result can be extended in a natural way to directed networks by noting that all vertices that are only targets of spillovers, never the origin of spillovers, will receive no allocation in equilibrium.

exists an efficient allocation, $\mathbf{e}^* \in \mathcal{E}^*(G)$*, that:*

1. *Allocates 0 to* u

2. *Allocates 1 to* v

If v *has two neighbors of degree 1, so there exist* $t, u, v \in V(G)$ *such that* $N(t) = N(u) = \{v\}$*, then the above conditions are true for every efficient allocation in* $\mathcal{E}^*(G)$*.*

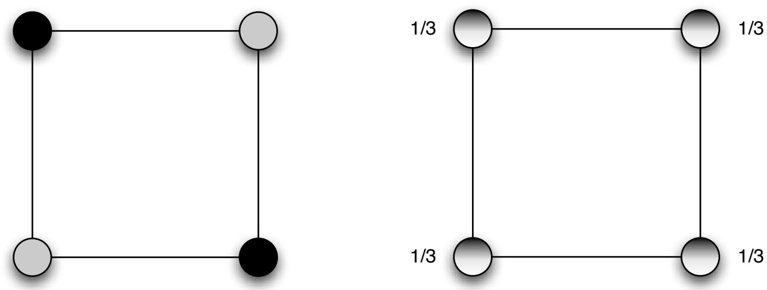
Proof of Theorem 2.5.2:

Consider some efficient allocation $\overline{\mathbf{e}}$ over G. We further consider two vertices $u, v \in V(G)$ with $N(u) = \{v\}$. Note that $\overline{e}_u, \overline{e}_v \leq 1$, since it is never efficient to allocate more than 1 to a vertex. Since u is a vertex of degree 1, and satisfies the targeting problem, $\overline{e}_u + \overline{e}_v = 1$.

Since u is of degree 1, the set of allocations to vertices other than u, $V(G) \setminus \{u\}$, constitutes an allocation that satisfies the targeting condition over $G \setminus \{u\}$. Formally, if $\overline{\mathbf{e}}$ is efficient on G, then the allocation $\overline{\mathbf{e}}^u$ defined by $\overline{e}^u_j = \overline{e}_j$ for $j \in V(G) \setminus \{u\}$, satisfies the targeting constraint, $\sum_{j \in N(i) \cup i} \overline{e}^u_j \geq 1$ for all $j \in V(G) \setminus \{u\}$. Consider an allocation, $\mathbf{e}^*$ over G with $e^*_u = 0$ and $e^*_v = 1$. It is clear that $\mathbf{e}^*$ is as efficient as $\overline{\mathbf{e}}$ over G and satisfies the targeting constraint for u. To see that $\mathbf{e}^*$ satisfies the targeting constraint over the whole network, note that $e^{*u}_j \geq \overline{e}^u_j$ for all $j \in V(G) \setminus \{u\}$.

Now consider two vertices of degree 1 with a common neighbor. Let $t, u, v \in V(G)$ satisfy $N(t) = N(u) = \{v\}$. Assume, without loss of generality, that there exists an allocation $\mathbf{e}'$ with $e'_u > 0$. Since $e'_u + e'_v \geq 1$ and $e'_t + e'_v \geq 1$, we know that $e'_t + e'_u + e'_v \geq 1 + e'_u$. This implies that an allocation $\mathbf{e}^*$ with $e^*_t = e^*_u = 0$ and $e^*_v = 1$ is strictly more efficient than $\mathbf{e}'$. The efficiency of $\mathbf{e}^*$ follows from the arguments above. $\square$

The theorem shows that vertices of degree 1 are likely to experience inequalities when it comes to receiving. We might be interested in scenarios where the allocator will be induced to give a more equal distribution of goods. There are certainly any number cases where "dividing the allocation" yields more efficient outcomes. In figure 2.17, the allocator pays a total of 2 in the efficient indivisible allocation (figure 2.17(a)) and a total of $\frac{4}{3}$ in the divisible case (figure 2.17(b)):



(a) Indivisible Allocation (b) Divisible Allocation

Figure 2.17: Efficiency Over Indivisible and Divisible Allocations

Figure 2.17(a) shows the efficient indivisible allocation, whereas figure 2.17(b) shows an efficient divisible allocation (where each vertex receives $\frac{1}{3}$). Clearly, the divisible case yields a more efficient allocation.

The next theorem derives necessary conditions for the efficient divisible allocation to yield strictly a more efficient outcome than the efficient indivisible allocation. Basically, the more efficient divisible allocation occurs when, in the indivisible case, at least two vertices receiving positive allocation share neighbors and are not linked to any vertices of degree 1 in the network.

Theorem 2.5.3. *If the efficient allocation over divisible goods is more efficient than that over indivisible goods, then there exist two vertices receiving positive allocation in an efficient indivisible solution that share neighbors and have no vertices of degree 1 in the spillover network. More precisely, if for any* $\mathbf{e} \in \mathcal{E}^*(G)$*, there exists some* $i \in V(G)$ *where* $e_i \notin \{0, 1\}$*, then for some minimum dominating set* I *and for some* $u, v \in I$ $(u \neq v)$*:*

1. $N(u) \cap N(v) \neq \emptyset$

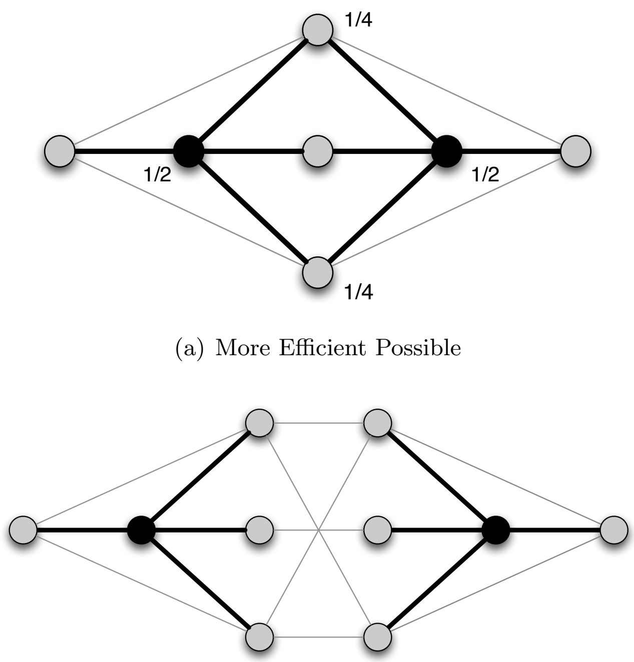
2. *If* $w \in N(u)$ *or* $w \in N(v)$*, then there exists some other* $r \in V(G)$ *such that* $w \in N(r)$ *(i.e. w is not a vertex of degree 1)*

Proof of Theorem 2.5.3:

Let the set of vertices receiving positive allocation in the indivisible allocation, I, have cardinality m. Assume the conditions do not hold, so we have a set $K \subsetneq I$ (possibly empty) containing vertices that satisfy condition 2 and not 1 (i.e. they share all neighbors with those vertices receiving positive allocation, but those vertices with which they share neighbors are connected to a vertex of degree 1). Consider the most efficient divisible solution, $\mathbf{x}$. For each $u \in I \setminus K$ (vertices that do not satisfy condition 2), allocate $1 - x_u$ for $x_u \geq 0$ (which represents all the possible allocations to the vertex). Notice, however, that each element of $I \setminus K$ is connected to a vertex of degree 1 since condition 2 fails. In the divisible allocation, these vertices of degree 1 must be allocated x_u in order to satisfy the targeting constraints. For each $v \in K$, let the sum of allocations to the neighbors of v be x_v in $\mathbf{x}$. Then, it follows that the allocation to v must be at least $1 - x_v$. But, then, in our allocation, we have:

$$\sum_{i \in V} x_i \geq \sum_{K} (x_v + 1 - x_v) + \sum_{I \setminus K} (x_u + 1 - x_u) = m$$

This shows, that we can never do better than the efficient indivisible solution if conditions 1 or 2 fail to hold. $\square$



(a) More Efficient Possible

(b) More Efficient Impossible

Figure 2.18: Possibility of a More Efficient Divisible Allocation

Figure 2.18(a) shows the efficient allocation over indivisible goods of a graph where the we can find a more efficient divisible allocation (the small numbers), whereas figure 2.18(b) shows an efficient indivisible allocation that cannot be improved upon.

Chapter 3

Analyzing Randomized Experiments with Spillovers

Abstract

This paper develops a general inferential framework for causal identification in randomized experiments in the presence of spillovers. Existing approaches focus on models of the underlying stochastic process governing spillovers or a priori knowledge of exactly which units share spillovers. This paper shows that the researcher may identify causal quantities of interest, without such strong assumptions, by analyzing the experiment with respect to inclusion probabilities induced by increasingly strong "social distance" restrictions. The social distance approach characterizes a fully general framework for causal identification in the presence of spillovers. Necessary assumptions for causal identification, as well as quantities that can be feasibly estimated, are discussed in detail. Using this framework, this paper develops an estimation strategy for causal identification with spillovers using thin-plate regression splines (TPRS). Above all, this paper demonstrates that the analysis of experiments in the presence of spillovers is feasible under reasonable, intuitive assumptions.

3.1 Introduction

Randomized experiments typically depend upon the strong assumption of no spillovers between units; that is, when one experimental unit receives the treatment, the effect of the treatment may not impact any other experimental unit through "spillovers." However, researchers are often interested in: 1) treatments that exhibit spatial and social network spillovers, such as information/advertising and vaccines; and 2) making empirical claims about spillovers, such as peer effects on the spread of information and the contagion effect in reducing illness through vaccines. This paper develops an inferential framework and a corresponding estimation strategy to deduce claims about randomized experiments in the presence of spillovers.[1]

Over the past few decades, the standard Rubin Causal Model (Rubin, 1974), which does not allow for spillovers, has been systematically extended to include for "partial interference" (Halloran and Struchiner, 1991; Sobel, 2006; Rosenbaum, 2007; Hudgens and Halloran, 2008; Tchetgen Tchetgen and VanderWeele, 2010). In such a setting, subject are arranged in "blocks," where it is assumed that spillovers may occur within blocks but not across them. Typically, however, spillovers over a social network, as in the examples above, cannot be neatly partitioned into blocks, and it is not known a priori which units share spillovers. This paper extends the partial interference framework and shows that, even in this more complicated setting, experimental inferences are feasible under reasonable assumptions. Furthermore, the partial interference framework and classic Rubin Causal Model emerge as special cases of this general framework.

Intuitively, one expects peer effects over a social network to weaken as the frequency or likelihood of interaction diminishes between two individuals. Similarly, one expects the extent of spatial spillovers to diminish as two experimental units are placed geographically further apart. This insight serves as the basis for the inferential framework in this paper. In

[1] I would like to thank Peter Aronow, Jake Bowers, Albert Fang, Mark Fredrickson, Donald Green, Dominik Hangartner, Macartan Humphreys, Ryan Moore, David Nickerson, Cyrus Samii and participants at the European Political Science Association Annual Meeting 2013, the 6th Annual Meeting of the Political Networks Section of the American Political Science Association, the American Political Science Association Annual Meeting and the Columbia University Methods Workshop for useful comments and advice. I am deeply grateful to Alexander Coppock, with whom a related paper has been co-authored, for continuing engagement and conversation on the topics discussed in this paper. All errors are my own.

particular, spillovers are analyzed with respect to a "social distance," such as frequency of interaction or geographic distance. As experimental units placed further and further apart according to the social distance are analyzed, estimators of a quantity of interest which condition upon the probability of assignment to appropriate treatment conditions, resulting from known randomization probabilities, converge in expectation to the true value. In short, by making successively more stringent distance restrictions, biases in the estimator resulting from spillovers may be removed. It can be shown that this social distance formulation is fully general; that is, any information the analyst can bring to bear on the problem of spillovers can be reformulated as a social distance.

Any number of social distance measures may suffice for the analysis. The social distance measure need only satisfy the *non-cascading assumption* – that spillovers eventually die out between experimental units placed further and further apart according to the social distance measure. Practically speaking, however, randomized experiments will rarely have the power, and experimental units will rarely dispersed enough, for this technique to fully converge to true quantity of interest. Accordingly, stronger conditions on the social distance measure are required for the analysis to be practicable.

3.1.1 Overview and Contribution

Quantities of interest are formed from *reference assignments*, idealized treatment assignment vectors such that one is guaranteed to see the potential outcome of interest. For instance, even in the context of spillovers, one is guaranteed to observe a fully untreated outcome for a particular unit under the treatment assignment vector that assigns no unit in the sample to receive the treatment (since this guarantees that there is no danger of spillovers from any other unit). Comparing observed outcomes, and estimated quantities, to these idealized quantities of interest yields a robust statistical inferential framework that allows for the analysis of bias and efficiency.

In order to implement these insights, the paper discusses an estimation strategy, with associated assumptions, to provide a fully practicable framework. In particular, a convexity assumption is imposed; that is, the expected magnitude of the bias is a convex (and decreasing) function of the social distance measure. It is argued that most intuitive social distance measures satisfy this assumption, and this assumption allows the researcher to bound the magnitude of the bias. Furthermore, estimators using thin-plate regression splines (TPRS) are developed to provide reasonably efficient estimation of quantities of interest, providing superior performance to standard inverse-probability weighted (IPW) estimators in a similar setting.

This paper makes several contributions to statistical analysis in a randomized experimental setting with spillovers. Existing approaches focus on models of the underlying stochastic process governing spillovers or knowledge a priori of exactly which units share spillovers. This paper shows that the researcher may retrieve a reasonable estimate of the quantity of interest without these sorts of strong assumptions if she can find a suitable candidate for the social distance measure without selecting an underlying stochastic process. Furthermore, this paper develops a self-contained statistical inferential framework through which one can assess the bias and efficiency of various approaches to causal estimation under spillovers. Finally, this paper extends upon existing methods for consistent estimation of causal quantities of interest by proposing a Bayesian penalized spline estimator which provides superior performance in efficiency without generating serious biases in the process.

Section 2 discusses the related literature and motivates the subject of this paper. Section 3 derives the inferential framework from a formal perspective, focusing on defining reference assignments, treatment conditions, and quantities of interest. Section 4 derives a set-theoretic formulation of interference structures, and develops a mathematical construction of the social distance measure. Section 5 develops the estimation strategy and assesses the performance of the proposed Bayesian thin-plate regression spline technique through simulation. Section 6 concludes the paper.

3.2 Related Literature and Motivation

The assumption of "no interference between units" (Cox, 1958) is commonly made in experiments and is one component of the stable unit treatment value assumption, or SUTVA (Rubin, 1980). Non-interference, however, may not always hold between experimental units, as with treatments that are likely to spillover such as advertising or vaccination.

In the past decade, researchers have begun seriously investigating contexts involving spillovers through experimental methods. Quantities of interest from such studies include, for instance, the effects of vaccination on infection rates in surrounding areas, the effects of get-out-the-vote campaigns on neighbors' voting rates, and the effects of educational interventions for students on their classmates.[2]

Recent work on the topic has focused upon blocked double randomization designs, originally described by Halloran and Struchiner (1991), where spillovers may occur within blocks but not across them. Tchetgen Tchetgen and VanderWeele (2010) show that IPW estimation, conditional on the probability of an outcome of interest being observed, can be used for unbiased estimation if there is a priori knowledge of exactly which units share spillovers. The IPW estimator, in the context of known spillover structures, has also been considered in Chen et al. (2010), Gerber and Green (2012), and Aronow and Samii (2013). Unfortunately, spillovers over more complex spaces, such as social networks and geography, do not fit into this framework since the data do not form discernible blocks, and the researcher does not know a priori which experimental units share spillovers. Spillover structures cannot just be assumed and compared because, unlike classical regression techniques, there is no way to assess the "fit" of an assumed spillover structure and corresponding estimator to an underlying causal parameter of interest. This paper demonstrates that a more general approach

[2]Such studies have often been conducted through blocked designs or double randomization (e.g., Duflo and Saez (2003), Giné and Mansuri (2011), Sinclair et al. (2012)). This yields three types of experimental groups, consisting of: (a) those units that are directly treated; (b) those units that are not directly treated but are located in clusters where units were treated (thus experiencing spillovers); and (c) those units that are not directly treated and are in clusters where no other unit was treated (serving as a control group). The average indirect effect is estimated as the average outcome in group (b) minus the average outcome in group (c). The statistical foundations of estimation via double randomization are discussed in detail in Hudgens and Halloran (2008).

that conditions upon probabilities of assignment at various distance restrictions can fit into a proper inferential framework when the spillover structure is unknown. The classic Rubin Causal Model and blocked randomization emerge as special cases of the framework.

The framework in this paper allows for inferences even when the researcher does not know the structure of spillovers and constitutes a generalization of the classic Rubin Causal Model (Rubin, 1974), as well as frameworks that require a known spillover structure such as Hudgens and Halloran (2008). In particular, following the insights of Rubin (1990) and Sobel (2006), potential outcomes are defined as a function of the vector of treatment assignments resulting from the randomization. This idea is used to define "reference assignments," which are assignment vectors that guarantee the potential outcome of interest will be observed (Rosenbaum, 2007). For example, one is guaranteed to see the fully untreated potential outcome for each unit under the assignment vector where no unit receives treatment. Following Sobel (2006), the quantity of interest is defined as a function of the number of reference assignments. This approach defines quantities of interest separately from assumptions about spillovers, thereby generating a self-contained inferential framework which allows for a characterization of bias and efficiency.

This framework does not require any parametric or model-based assumptions about the stochastic process governing spillovers. While this provides a more objective approach to estimating quantities of interest, it should be noted that estimation based upon the IPW estimator tends to be quite inefficient (Basu, 1971). For this reason, section 5 develops an estimation strategy to increase efficiency by using thin-plate regression splines. From a model-based perspective, Bowers et al. (2013) provide a novel Fisherian approach to assessing the plausibility of various theory-driven models of spillovers.

The approach in this paper differs from the existing literature in three ways. First and foremost, this approach does not require the researcher to know the exact spillover structure or stipulate an underlying spillover process, greatly enhancing the objectivity of the estimation procedure. Second, this paper characterizes the class of quantities that can be reasonably estimated from randomized experiments under spillovers and provides a principled approach to constructing such quantities. Finally, while inverse-probability weighted estimators may provide consistent estimation, they are known to be highly inefficient. This paper develops a semi-parametric estimator using thin-plate regression splines to generate a nearly consistent estimation strategy with superior performance in terms of efficiency, providing more robust estimation of quantities of interest.

3.3 General Framework

This section develops a principled approach to constructing quantities of interest when treatment effects can spill over units in the sample population. In particular, following Rubin (1990) and Sobel (2006), general potential outcomes are defined as a function of each possible vector of treatment assignments. The quantities of interest, however, are defined with respect to reference assignments, idealized treatment assignment vectors over which one is guaranteed to see a potential outcome of interest. The notion of a reference assignment was developed by Rosenbaum (2007), which he referred to as a "uniformity trial."[3] One of the key innovations of this paper is to notice that a meaningful statistical inferential framework results from the comparison of estimators, constructed from observed treatment assignment vectors, to this idealized quantity, constructed from reference assignments.

In this section, it is shown that reasonable causal inferences can be made under intuitive assumptions, even if the researcher does not know the underlying stochastic form governing spillovers. In particular, estimators are defined with respect to restrictions on social distance between units. As more restrictive assumptions are made, these estimators, which condition on the probability of assignment, converge in expectation to the quantity of interest. This occurs when the social distance measure satisfies a non-cascading assumption; that is, the strength of spillovers between two units diminish as the social distance between them increases. The flexibility of the framework results from the fact the researcher need not choose

[3]Originally, Rosenbaum (2007) used the phrase "uniformity trial" to refer to a situation where treatment was withheld from every unit to guarantee observation of the outcome for an untreated unit without spillovers, but in principle this idea can be generalized to any underlying "reference assignment," not just the no treatment case.

one "true" social distance measure; rather, any social distance measure that satisfies the non-cascading assumption will suffice for analysis. In practice, however, more restrictive assumptions will have to be placed upon the social distance metric in order to make inferences about the amount of bias, as discussed in section 4.

The steps of the inferential framework are as follows:

1. Classify *reference assignments* that guarantee observation of potential outcomes of interest

2. Form *treatment conditions*, collections of reference assignments that are consistent with an exposure type of interest (e.g., direct, indirect)

3. Generate *quantities of interest* which are typically formed as the difference between two treatment conditions

4. Identify a desirable *social distance* measure

5. Iteratively calculate *estimators conditioning on probability of assignment* under more and more stringent social distance restrictions, which converge in expectation to the quantities of interest

The first 3 steps of this process are covered in this section. Section 4 discusses step 4, identifying a social distance measure. Finally, section 5 discusses the estimators and general estimation strategy.

Section 3.3.1 defines the types of quantities that can estimated in an experimental setting, which will be discussed in the remainder of the paper. Section 3.3.2 develops the ideas of references assignments and treatment rigorously, and section 3.3.3 defines the quantities of interest rigorously, using the ideas of reference assignments and treatment conditions.

3.3.1 Quantity of Interest: Average Exposure Effect

Once a treatment may spill over to other units in the sample, the researcher must make difficult choices about what she intends to estimate. Not every quantity of interest will be estimable using experimental approaches. To understand this point, it is useful to make a distinction between an *average equilibrium effect* and an *average exposure effect*.

An average equilibrium effect measures the population average effect of the treatment if the treatment regime was "fully implemented" in the population. Suppose the researcher is interested in the average effect of a get-out-the-vote campaign encouraging women to vote on female voter turnout, as in Giné and Mansuri (2011). Suppose further that the researcher defines full implementation of the campaign as targeting each household in the village. With no spillovers, the average treatment effect estimated from a sufficiently large experiment, where a randomly selected subset of the population is given the treatment, is a good estimate of the average treatment effect if the treatment is fully implemented. However, this is not true if the treatment, the get-out-the-vote campaign, has spillovers over the population. The quantity measuring the average effect of the information campaign under full implementation *cannot* be measured using an experiment without strong assumptions. To see this, notice that any experiment would only treat some subset of the village. However, the potential outcome of any unit under spillovers is necessarily a function of the treatment status of every other unit in the population and thus the average outcome when every unit is treated cannot generally be estimated from an experiment which only treats a subset of the population. In such a scenario, certain non-experimental techniques, such as interrupted time series or pre-post designs, may be preferred.

An average exposure effect measures the population average effect of a certain type of "exposure" isolated for each unit. For instance, one may be interested in the average effect of a unit directly receiving the campaign and experiencing no other spillovers, or the average effect of having exactly one adjacent household receiving the campaign when untreated, or the average effect of having at least two adjacent households receiving the campaign when untreated. This sort of average isolated effect can be estimated well by experimental techniques under certain conditions, which are detailed below. The reason for measuring average exposure effects is that it allows the researcher to develop a more nuanced understanding of spillover effects in the population. For instance, the researcher can estimate a direct treatment effect, the effect of having a single neighbor treated when untreated, the effect of having two neighbors treated when untreated, and so on, in order to map out the strength of the direct effect and spillovers in the population. These sorts of quantities are particularly useful when the researcher must estimate treatment effects over an entire population when only being allowed to treat a portion of it; these nuanced measures can provide data on how to most effectively target a population given a fixed number of units to treat or a fixed budget. It will be assumed throughout the rest of this paper that the quantity of interest follows the form of an average exposure effect.

The average exposure effects estimated in this paper are similar to those described by Sobel (2006), with one difference. Sobel (2006) defines quantities of interest as averages over possible treatment assignments in a particular randomization scheme. By contrast, this paper averages over reference assignments, idealized treatment assignment vectors over which one is guaranteed to see a potential outcome of interest. This idealized form for the quantity of interest is necessary because it allows for an analysis of bias and consistency and its relationship to social distance.

3.3.2 Treatment Conditions, Reference Assignments, and Potential Outcomes

Let $\mathcal{R}$ be a space associated with a "social distance" or quasi-metric $\rho : \mathcal{R} \times \mathcal{R} \to [0, \infty]$ (Wilson (1931)).[4] In other words, ρ is a like a traditional metric, but it may be the case that $\rho(x, y) \neq \rho(y, x)$, which is common is many social settings.[5] Consider a set of experimental units, $\Omega = \{1, \ldots, N\}$, drawn from $\mathcal{R}$ such that for any $x, y \in \Omega$, one may associate $\rho(x, y)$. Intuitively, this means that the experimental units are endowed with social distance which will be used to characterize how far spillovers may travel between the units. The implications

[4] The quasi-metric ρ is still non-negative and satisfies the triangle inequality. Given $x, y, z \in \mathcal{R}$:

- $\rho(x, y) \geq 0$ with $\rho(x, x) = 0$
- $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$

[5] This is seen in directed social networks. Intuitively, if we think of distance as the likelihood of interaction for spreading a disease, the likelihood of spreading from x to y may not be the same as spreading from y to x.

of social distance will be discussed in section 4.

Formally, each unit i is assigned to one of t treatments, $d_i \in \{0, 1, \ldots, t-1\}$,[6] and a treatment assignment vector over the entire population will be denoted as the N-tuple $\mathbf{d} \in \mathbf{D}$, where $\mathbf{D} = \{0, 1, \ldots, t-1\}^N$.

Let $y : \mathbf{D} \to \mathbb{R}^N$ be a real vector-valued potential outcome function over the treatment assignment draw, where $y_i(\mathbf{d})$ denotes the outcome for subject i associated with a particular treatment assignment vector, $\mathbf{d}$. In studies with spillovers, the quantity of interest is typically the expected difference in the population between two *treatment conditions*. In the presence of spillovers, a treatment condition is any subset of treatment assignment vectors defined for each unit that necessarily satisfies the condition of interest (e.g., spillover from a single neighboring unit).

In order to construct the treatment conditions of interest, it is first necessary to properly define *reference assignments*, which are treatment assignment vectors that will guarantee that the outcome of interest is observed. These will serve as the rudiments of the treatment conditions of interest and will assure that the construction of quantities of interest are not dependent upon modeling assumptions. The following two definitions formally explicate these ideas:

Definition 3.3.1 (Reference Assignments and Exposure Conditions)**.** *A subset of entire space of treatment assignment vectors, $\mathbf{S}_i \subseteq \{0, \ldots, t-1\}^N$, that satisfies some logical condition such that each element may admit a different observed outcome is defined as an **exposure condition** for unit i. That is, given $\mathbf{d}, \mathbf{d}' \in \mathbf{S}_i$, it follows that $y_i(\mathbf{d}) \not\equiv y_i(\mathbf{d}')$. Each $\mathbf{d} \in \mathbf{S}_i$ is referred to as a **reference assignment** for the exposure condition.*

Definition 3.3.2 (Treatment Conditions)**.** *A **treatment condition** is defined by a collection of exposure conditions, one for each unit, $\mathcal{S} = \{\mathbf{S}_1, \ldots, \mathbf{S}_N\}$ where $\mathbf{S}_i \subseteq \{0, \ldots, t-1\}^N$ is a subset of treatment assignment vectors (possibly empty) for each $i \in \Omega$ that satisfies some*

[6] Throughout the paper $d_i = 0$ will be used as a control condition.

logical condition. For clarity of exposition, $\mathbf{S}_i \in \mathcal{S}$ *will be referred to as the treatment condition with respect to* i*.*

Intuitively, reference assignments are those assignment vectors for which the potential outcome of interest is guaranteed to be observed. As an example, the untreated condition will typically be written as $\mathcal{S}_0 = \{\{\mathbf{0}\}, \ldots, \{\mathbf{0}\}\}$, where the reference assignment for the untreated condition for each unit is $\mathbf{0}$, the assignment vector where each unit receives $d_i = 0$. One is guaranteed to observe the potential outcome in the fully untreated condition for a particular unit when every experimental unit, i, is assigned to $d_i = 0$.

An Example

In order to make the formal framework easier to comprehend, the formal details will be presented alongside a toy problem. Consider the following problem:

> *Over a population of 50 individuals, 10 individuals are randomly selected to receive the treatment, an advertisement to buy some product.*[7] *The researcher is interested in the following questions: What is the average increase in expenditure on the product if the individual is directly targeted by the advertisement? What is the average increase in expenditure on the product if a friend is directly targeted by the advertisement while the individual is/is not directly targeted by the advertisement?*

The goal is to isolate the usual direct treatment effect, the average effect of having a single treated neighbor treated while untreated, and the average effect of having a single treated neighbor while treated. The *treatment* is the advertisement. The four *treatment conditions* of interest are: 1) the direct exposure condition–the direct effect of being treated without any other interference; 2) the indirect exposure condition–the spillover effect from a neighboring treated friend on an untreated individual; 3) the joint exposure effect–the combined of effect of receiving direct treatment and having a friend that is directly treated; and 4) the untreated

[7]This is a problem that has been studied extensively in the game-theoretic literature under the topic of "word-of-mouth advertising." See Galeotti and Goyal (2009) for an overview.

condition–an untreated individual experiencing no spillovers. The related quantities of interest are the *direct treatment effect*, the average difference between the direct and untreated conditions, the *indirect exposure effect*, the average difference between the indirect exposure and untreated conditions, and the *joint exposure effect*, the average difference between the joint exposure and untreated conditions.

Estimation problems involving spillovers are often best visualized over a network or a mathematical graph. Causal effects are estimated for a population of N individuals, the *units* of our analysis. The sample space is $\Omega = \{1, \ldots, N\}$ with $N = 50$. An *edge* is formed between two individuals in the network if they are friends, where the set of such edges is denoted by $\mathcal{E} \subseteq \Omega \times \Omega$.[8] $\mathcal{G}$ is uniquely defined by the pair $\mathcal{G} = (\Omega, \mathcal{E})$.

Figure 3.1 is a representation of the problem with 50 individuals. The individuals are represented as *nodes* of the graph, and an edge exists between any two nodes if the corresponding individuals are friends. The "neighbors" of an individual, i, are those individuals who are friends of i. As shown in figure 3.1, the neighbors of i, $\zeta(i)$, are described by those nodes that share an edge with i. The *degree* of node i, $\delta_i = |\zeta(i)|$, is the number of neighbors for node i in the graph.

In this experiment, the treatment is binary, $d_i \in \{0, 1\}$. The treatment conditions of interest are the direct treatment condition ($\mathcal{S}_1$), the indirect exposure condition ($\mathcal{S}_{01}$), the joint exposure condition ($\mathcal{S}_{11}$) and the untreated condition ($\mathcal{S}_0$). Let $\mathbf{d}_j = \{\mathbf{d} | d_j = 1, d_k = 0, k \neq j\}$ denote a vector of treatment assignments where unit j is assigned to treatment and all other units have been assigned to no treatment. Furthermore, let $\mathbf{0}$ denote the vector of treatment assignments where each unit is assigned to no treatment.

Figure 3.2 describes possible reference assignments. Figure 3.2(a) displays the only reference assignment for the direct treatment condition for unit i, the assignment vector that assigns i to be directly treated while no other unit is treated. This implies that one is guaranteed to see the outcome associated with unit i being directly treated without any other

[8] Although our discussion assumes "undirected edges," i.e. an edge between i and j implies and edge between j and i, this framework is applicable to "directed networks" as well, where this property need not be true. Also note that this implies that each undirected edge is counted as two reciprocal directed edges in $\mathcal{E}$.

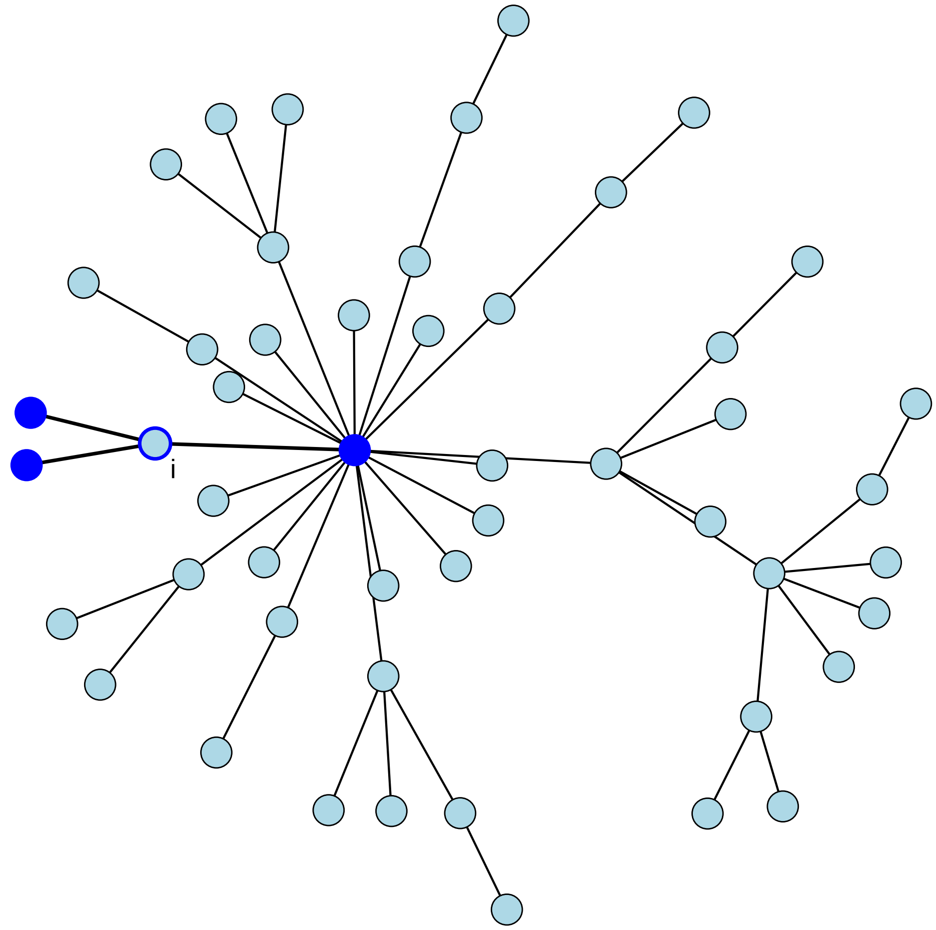
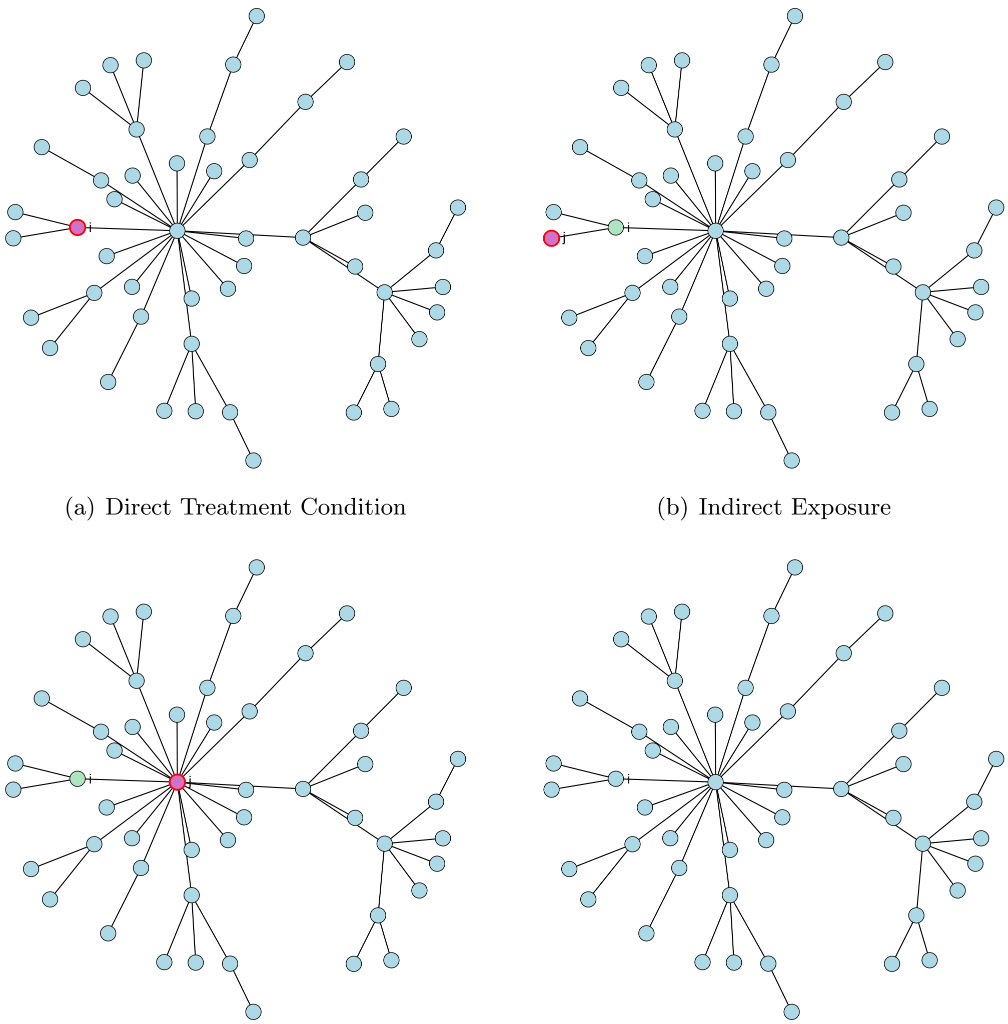


Figure 3.1: An Example of a Network or Graph with 50 Nodes

Figure 3.1 shows a graph or network, $\mathcal{G}$. The set of individuals is denoted by the nodes in the network. The edges between two nodes denote farmers with adjacent plots of land. The set of neighbors of i, $\zeta(i)$, are shown as the nodes shaded in a darker blue. The degree of i is $\delta_i = |\zeta(i)| = 3$. The network is drawn to have a "scale-free" distribution, which means that degree distribution for the nodes in the network follow an inverse power law, a pattern typically observed in social networks. This network is generated from a Barabási-Albert model with $m = 1$ (Albert and Barabási, 2002).

interference under the treatment assignment vector that assign i to treatment and all other units to no treatment. Figures 3.2(b) and 3.2(c) display two possible reference assignment for treatment condition of indirect exposure from having exactly one neighbor treated. There are a total of three such reference assignments, corresponding to treatment assignment vectors where exactly one neighbor of i is treated and no other units receive treatment. These imply that one is able to observe the outcome of having exactly one neighbor treated under the

Figure 3.2: Various Reference Assignments



(a) Direct Treatment Condition

(b) Indirect Exposure

(c) Indirect Exposure

(d) Untreated

Figure 3.2 describes possible reference assignments. Figure 3.2(a) displays the only reference assignment for the direct treatment condition for unit i, the assignment vector that assigns i to be directly treated while no other unit is treated. Figures 3.2(b) and 3.2(c) display two possible reference assignment for treatment condition of having exactly one neighbor treated. There are a total of three such reference assignments, corresponding to treatment assignment vectors where exactly one neighbor of i is treated and no other units receive treatment. Figure 3.2(d) displays the only reference assignment for the untreated treatment condition for unit i, the assignment vector that assigns no unit to the treatment condition.

treatment assignment vector that assigns one such neighbor to treatment and every other unit to no treatment. Notice, however, that there are multiple such reference assignments; this is because the outcome is potentially different based on which neighbor is treated. Finally, as before, the only reference assignment for the untreated condition is the assignment vector that assigns no unit to the treatment, as shown in figure 3.2(d).

The four treatment conditions of interest can be written as follows:

$$\mathcal{S}_1 = \{\{\mathbf{d}_1\}, \ldots, \{\mathbf{d}_N\}\} \tag{3.3.1}$$

$$\mathcal{S}_{01} = \{\{\mathbf{d}_j, j \in \zeta(1)\}, \ldots \{\mathbf{d}_j, j \in \zeta(N)\}\}$$

$$\mathcal{S}_{11} = \{\{\mathbf{d}_1 + \mathbf{d}_j, j \in \zeta(1)\}, \ldots \{\mathbf{d}_N + \mathbf{d}_j, j \in \zeta(N)\}\}$$

$$\mathcal{S}_0 = \{\{\mathbf{0}\}, \ldots, \{\mathbf{0}\}\}$$

The potential outcomes in $\mathcal{S}_1$ are of the form $y_i(\mathbf{d}_i)$, the potential outcomes in $\mathcal{S}_{01}$ are of the form $y_i(\mathbf{d}_j)$ if $j \in \zeta(i)$, he potential outcomes in $\mathcal{S}_{11}$ are of the form $y_i(\mathbf{d}_i + \mathbf{d}_j)$ if $j \in \zeta(i)$, and the potential outcomes in $\mathcal{S}_0$ are of the form $y_i(\mathbf{0})$.

3.3.3 Quantity of Interest

The quantity of interest is defined over treatment conditions. In particular, a quantity will typically involve the (weighted) average of the difference in potential outcomes between the treatment conditions over all of the units, defined through the reference assignments. The notation for a quantity of interest is made more complex by the fact that each unit may have a different number of potential outcomes within and across treatment conditions (as in the example for the indirect and untreated treatment conditions). For a fixed set of units, Ω, the quantity of interest is defined as follows:[9]

Definition 3.3.3. *Let a **quantity of interest**, $\tau : \mathcal{D} \times \mathcal{D} \rightarrow \mathbb{R}$, be defined over Ω. Then,*

$$\tau(\mathcal{S}, \mathcal{S}') = \frac{1}{\sum_{i \in \Omega} |\mathbf{S}_i||\mathbf{S}'_i| w_i} \sum_{i \in \Omega} \sum_{\mathbf{d} \in \mathbf{S}_i} \sum_{\mathbf{d}' \in \mathbf{S}'_i} w_i y_i(\mathbf{d}) - w_i y_i(\mathbf{d}')$$

[9] i.e., $\mathcal{S}$ and $\mathcal{S}'$ have non-empty components for the corresponding units in Ω

$$= \frac{1}{\sum_{i\in\Omega}|\mathbf{S}_i||\mathbf{S}'_i|w_i}\sum_{i\in\Omega}\sum_{\mathbf{d}\in\mathbf{S}_i}|\mathbf{S}'_i|w_i y_i(\mathbf{d}) - \frac{1}{\sum_{i\in\Omega}|\mathbf{S}_i||\mathbf{S}'_i|w_i}\sum_{i\in\Omega}\sum_{\mathbf{d}'\in\mathbf{S}'_i}|\mathbf{S}_i|w_i y_i(\mathbf{d}')$$

where w_i are unit-specific weights.

It will be useful at times to deal with the two terms in the final equation separately.

Definition 3.3.4. *The* ***weighted expected value under the treatment condition*** $\mathcal{S}$*, with respect to another treatment condition* $\mathcal{S}'$*, is given by:*

$$\overline{y(\mathcal{S},\mathcal{S}'_i)}_w = \frac{1}{\sum_{i\in\Omega}|\mathbf{S}_i||\mathbf{S}'_i|w_i}\sum_{i\in\Omega}\sum_{\mathbf{d}\in\mathbf{S}_i}|\mathbf{S}'_i|w_i y_i(\mathbf{d})$$

Thus, the quantity of interest may be written as $\tau(\mathcal{S},\mathcal{S}') = \overline{y(\mathcal{S},\mathcal{S}')}_w - \overline{y(\mathcal{S}',\mathcal{S})}_w$. Note that the definition of a quantity of interest does not depend on non-interference assumptions in this setting. Non-interference assumptions, which are necessary for estimation, are made independently of the quantity of interest.

Application to the Example

In order to define a more intuitive notation similar to Hudgens and Halloran (2008), let $y_i(\mathbf{d}_i) \equiv y_i(1)$, $y_i(\mathbf{d}_j) \equiv y_{ij}(0,1)$, $y_i(\mathbf{d}_i + \mathbf{d}_j) \equiv y_{ij}(1,1)$, and $y_i(\mathbf{0}) \equiv y_i(0)$, where $\mathbf{d}_j = \{\mathbf{d}|d_j = 1, d_k = 0, k \neq j\}$. The advantage of this notation is that it become clear that there is a potential outcome for each pair of nodes i, j, $j \in \zeta(i)$. It follows that the quantities of interest, the direct treatment effect, $\tau(\mathcal{S}_1,\mathcal{S}_0)$, the indirect exposure effect, $\tau(\mathcal{S}_{01},\mathcal{S}_0)$, and the joint exposure effect, $\tau(\mathcal{S}_{11},\mathcal{S}_0)$ can be written as:

$$\tau(\mathcal{S}_1,\mathcal{S}_0) = \frac{1}{N}\sum_{i=1}^{N} y_i(1) - y_i(0) = \frac{1}{N}\sum_{i=1}^{N} y_i(1) - \frac{1}{N}\sum_{i=1}^{N} y_i(0) \tag{3.3.2}$$

$$\tau(\mathcal{S}_{01},\mathcal{S}_0) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\delta_i}\sum_{j\in\nu(i)} y_{ij}(0,1) - y_i(0) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\delta_i}\sum_{j\in\nu(i)} y_{ij}(0,1) - \frac{1}{N}\sum_{i=1}^{N} y_i(0) \tag{3.3.3}$$

$$\tau(\mathcal{S}_{11}, \mathcal{S}_0) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\delta_i}\sum_{j\in\nu(i)} y_{ij}(1,1) - y_i(0) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\delta_i}\sum_{j\in\nu(i)} y_{ij}(1,1) - \frac{1}{N}\sum_{i=1}^{N} y_i(0) \quad (3.3.4)$$

The average direct effect, $\tau(\mathcal{S}_1, \mathcal{S}_0)$ is the analogue of the average treatment effect in a classical setup. In the indirect and joint exposure conditions for some unit i, there is a separate potential outcome corresponding to each unique neighbor j, while in the untreated condition there is one potential outcome for unit i. Thus, the number of potential outcomes (reference assignments) in the indirect and joint exposure conditions for unit i is equal to the degree of i, implying that each unit has a different number of potential outcomes within and across treatment conditions. In order to create an intuitive quantity, where each unit receives "equal" importance, the difference between the indirect/joint condition and the untreated condition is normalized by the degree of the unit. Note that $|\mathbf{S}_{1i}| = |\mathbf{S}_{0i}| = 1$ and $|\mathbf{S}_{01i}| = |\mathbf{S}_{11i}| = \delta_i$. In $\tau(\mathcal{S}_1, \mathcal{S}_0)$, $w_i = 1$, and in $\tau(\mathcal{S}_{01}, \mathcal{S}_0)$ and $\tau(\mathcal{S}_{11}, \mathcal{S}_0)$, $w_i = \frac{1}{\delta_i}$. The calculations follow from the fact that:

$$\sum_{i\in\Omega} |\mathbf{S}_{1i}||\mathbf{S}_{0i}| * 1 = \sum_{i\in\Omega} |\mathbf{S}_{01i}||\mathbf{S}_{0i}| * \frac{1}{\delta_i} = \sum_{i\in\Omega} |\mathbf{S}_{11i}||\mathbf{S}_{0i}| * \frac{1}{\delta_i} = N$$

3.4 Non-Interference Partitions and Social Distance

The previous section constructed quantities of interest without resorting to modeling assumptions. After one has determined reference assignment and quantities interest, the goal is to understand when these quantities can be isolated with little or no interference from other spillovers. This section describes a general approach to isolating these quantities. The construction in this section will follow two steps. First, the notion of non-interference will be precisely defined from a mathematical and set-theoretic point of view. Second, the section develops the idea of a "social distance" measure and its relationship to non-interference, as well as the generality of the framework. As the researcher considers the set of units that are further and further apart according to this social distance measure, biases due to complex spillovers are reduced. Thus, considering units subject to a sufficiently stringent social distance constraint allows for estimation of quantities of interest with lower levels of bias.

Section 4.1 describes the idea of interference from a set-theoretic perspective. Section 4.2 defines the social distance framework and describes how it applies to the toy problem. Section 4.3 discusses how the chosen distance restrictions affect the sample chosen for the quantity of interest; in particular, it is shown that considering more stringent distance restrictions also decreases the number of admissible units considered for the quantity of the interest. Finally, section 4.4 discusses the generality of the social distance approach, namely that information brought to bear on the problem can be reformulated in terms of a social distance measure.

3.4.1 Non-Interference Partitions

Non-interference assumptions entail partitioning the set of assignment vectors for which one believes the potential outcome of interest for unit i will be observed. A *non-interference partition* is a partition of the space of assignment vectors where each element of the partition is comprised of the treatment assignment vectors over which one observes the same potential outcome as the reference assignment. Thus, in this framework, non-interference assumptions entail making a guess at the (unknown) non-interference partition. For instance, under the standard non-interference assumption made in most experiments, the set of assignment vectors $\{\mathbf{d} \mid d_i = 1\}$ yields the same potential outcome as $\mathbf{d}_i$, the vector that assigns i to treatment and all other units to no treatment.

Each distinct potential outcome in the quantity of interest must be separately observable. This will imply that two reference assignments in the relevant treatment conditions for i, $\mathbf{S}_i$ and $\mathbf{S}_i'$, have non-overlapping non-interference partitions (due to the fact that they yield distinct potential outcomes).

Definition 3.4.1 (Non-Interference)**.** *Let Q^i be a **non-interference partition**. For ease of exposition, denote the element of the partition containing $\mathbf{d}$ as $Q^i(\mathbf{d})$. The following three conditions always hold for non-interference partitions:*

1. *Q^i is a partition of the space of assignment vectors,* $\mathbf{D}$.

2. *Let* $\mathbf{d}^* \in Q^i(\mathbf{d})$. *Then,* $y_i(\mathbf{d}) = y_i(\mathbf{d}^*)$.

3. *For* $\mathbf{d}, \mathbf{d}' \in \mathbf{S}_i \cup \mathbf{S}_i'$, $Q^i(\mathbf{d}) \cap Q^i(\mathbf{d}') = \emptyset$.

An experiment is a design where probability of observing each treatment assignment vector is known. The associated measure-theoretic definition of an experiment is given below.

Definition 3.4.2. *An **experiment** is defined by the triple* $(\mathbf{D}, \mathcal{F}, P)$, *where* $\mathbf{D} = \{0, \ldots, t-1\}^N$ *is a space of possible assignment draws, and* P *is a probability measure with* $P(\emptyset) = 0$ *and* $P(\mathbf{D}) = 1$. $\mathcal{F}$ *is the usual* σ*-algebra formed by unions, intersections, and complements from the elements of* $\mathbf{D}$.

The probability that an outcome $y_i(\mathbf{d})$ is observed under non-interference partition Q^i is just the probability measure of the partition containing the reference assignment $\mathbf{d}$:

$$\pi_{y_i(\mathbf{d})} = P(Q^i(\mathbf{d})) \tag{3.4.1}$$

In order for a quantity of interest to be *estimable*, each potential outcome in the quantity must be observable with positive probability,[10] and at least one unit in each treatment condition must be observed.

Definition 3.4.3 (Estimability). *Let* Q^i *denote the non-interference condition for unit* i. *A quantity,* τ, *over the units* Ω *is said to be **estimable** if:*

1. $\pi_{y_i(\mathbf{d})} > 0$ *for all* $\mathbf{d} \in \mathbf{S}_i \cup \mathbf{S}_i'$ *and all* $i \in \Omega$

2. *For each* $\mathbf{d} \in \mathbf{D}$, $P(\{\mathbf{d}\}) > 0$ *implies that there exists* $\mathbf{d}^i \in \mathbf{S}_i$, $\mathbf{d}^j \in \mathbf{S}_j'$ *such that* $\mathbf{d} \in Q^i(\mathbf{d}^i) \cap Q^j(\mathbf{d}^j)$ *for* $i, j \in \Omega$

[10] A more precise statement would be that for the quantity to be estimated without distributional assumptions, each potential outcome must be observed with positive probability. This will have implications for the consistency of the estimate under the Bayesian approach described below. See Rubin (1978) for more on this point.

To understand this framework, consider a standard experimental setup with binary treatment status and N units in the population. In this setting each unit, i has treatment assignment $d_i \in \{0,1\}$. Intuitively, when isolating the effect of $d_i = 1$ or $d_i = 0$, one assumes that all other units are not given any other treatment. Thus, the "treated potential outcome" for i, generally denoted as $y_i(1)$, can be written as $y_i(\mathbf{d}_i)$ where the reference assignment $\mathbf{d}_i = \{\mathbf{d}|d_i = 1, d_j = 0, j \neq i\}$ denotes an assignment draw where unit i has been assigned to treatment and all other units have been assigned to no treatment. Similarly, the "untreated potential outcome" for unit i, generally denoted as $y_i(0)$ may be written as $y_i(\mathbf{0})$, where $\mathbf{0}$ is the assignment vector where each unit in the population is assigned to no treatment. The two treatment conditions of interest are the isolated effect of treatment on a unit and the untreated condition for a unit, i.e. $\mathcal{S} = \{\{\mathbf{d}_1\}, \dots, \{\mathbf{d}_N\}\}$ and $\mathcal{S}' = \{\{\mathbf{0}\}, \dots, \{\mathbf{0}\}\}$. Averaging over all units, the usual quantity of interest can then be written as:

$$\tau(\mathcal{S}, \mathcal{S}') = \frac{1}{N}\sum_{i=1}^{N} y_i(\mathbf{d}_i) - y_i(\mathbf{0}) \tag{3.4.2}$$

This quantity, the standard direct treatment effect, requires no assumptions about non-interference. Independently of this experiment, $(\{0,1\}^N, \mathcal{F}, P)$ and associated quantities, the associated non-interference condition in this classical setting is defined as as $Q^i(\mathbf{d}_i) = \{\mathbf{d}|d_i = 1\}$ and $Q^i(\mathbf{0}) = \{\mathbf{d}|d_i = 0\}$. The probabilities of assignment to treated and untreated potential outcomes for i, generally denoted as $\pi_i(1)$ and $\pi_i(0)$, are defined as:

$$\pi_{y_i(\mathbf{d}_i)} = \pi_i^1 = P(Q^i(\mathbf{d}_i)) = P(\{\mathbf{d}|d_i = 1\}) \tag{3.4.3}$$

$$\pi_{y_i(\mathbf{0})} = \pi_i^0 = P(Q^i(\mathbf{0})) = P(\{\mathbf{d}|d_i = 0\})$$

While the usual complete non-interference condition has been stipulated, more complicated assumptions about spillovers may be made over the same data.

3.4.2 Social Distance and Convergence in Expectation

The statistical approach advocated in this paper is to stipulate successively "more restrictive" non-interference assumptions in terms of social distance. It will be shown that a class of estimators that condition upon the probability of assignment generally converge in expectation under this sequence of more restrictive assumptions. Since the estimators are broadly design unbiased/consistent, this technique produces causally identified estimates without direct knowledge of the interference structure.

Experimental units are endowed with social distance through the quasi-metric ρ defined in section 3.3.2. Interference with observing the potential outcome of unit i occurs if another treated unit, j, is "too close" to i. This suggests restrictions of the form $d_j = 0$ if $\rho(i, j) < c$, where c is some distance restriction imposed on the estimator. However, the relationship between c and the potential outcomes is unknown to the researcher. Consider an estimator $\widehat{\tau}$ that provides a consistent estimate of the quantity of interest, when conditioning upon known probabilities of assignment to respective treatment conditions. Let $\widehat{\tau(c)}$ denote the same estimator under the restriction implied by c; that is, the estimator is conditioned upon probabilities implied by the distance restriction c. The main mathematical result of this section is that it is generally possible to take an unbounded increasing sequence, $(c_k)_{k=1}^{\infty}$, such that $E\left(\widehat{\tau(c_k)}\right)$ forms a Cauchy sequence. This implies that there exists K such that $|E\left(\widehat{\tau(c_K)}\right) - E\left(\widehat{\tau(c_k)}\right)|$ is arbitrarily small for $k > K$.

The non-interference assumption implied by a distance restriction is stated formally below.

Definition 3.4.4. *Consider some constant $c \in [0, \infty]$ for each unit i with reference assignment $\mathbf{d} \in \mathbf{S}_i \cup \mathbf{S}'$. A **non-interference assumption induced by** c is a partition of assignment vectors that satisfies:*

$$\tilde{Q}_c^i(\mathbf{d}) = \{\mathbf{d} + \mathbf{d}' | d'_j = 0 \textit{ if } \rho(i, j) \leq c\},\ \mathbf{d}' \in \mathbf{D}$$

for each reference assignment $\mathbf{d} \in \mathbf{S}_i \cup \mathbf{S}'_i$.

It is important to note that the set of reference assignments and treatment conditions may not be identical over distance restrictions since some restrictions may not be stringent enough to permit observation. For instance, under the restriction $c = 0$, there are by definition no spillovers between units (as in the classical case). In this situation, the notion of indirect exposure does not exist so such a treatment condition is never observed for any unit. It is true, however, that any reference assignments and treatment conditions that exist at a particular distance restriction will necessarily exist at more stringent distance restrictions.

A second definition, which captures treatment conditions induced by the distance restriction c, will also be useful. The definition states that an observed treatment assignment vector, $\mathbf{d}^*$ satisfies a treatment condition induced by c for unit i, $\mathbf{S}_i(c)$ when there exists a reference assignment in the treatment condition for which $\mathbf{d}^*$ is contained in the non-interference assumption induced by c.

Definition 3.4.5. *A treatment assignment vector* $\mathbf{d}^*$ *is included in a treatment condition induced by c for unit i if there exists* $\mathbf{d} \in \mathbf{S}_i$ *such that* $\mathbf{d}^* \in \tilde{Q}^i_c(\mathbf{d})$. *This will be denoted by* $\mathbf{d}^* \in \mathbf{S}_i(c)$. *The treatment condition induced by c will be denoted as* $\mathcal{S}(c)$.

Using this definition, the estimator implied by a particular distance restriction, c, can be stated:

Definition 3.4.6. *For each reference assignment,* $\mathbf{d} \in \mathbf{S}_i$*, define:*

$$\pi_{y_i(\mathbf{d})}(c) = P(\tilde{Q}^i_c(\mathbf{d}))$$

For a given assignment vector $\mathbf{d}^*$*, the corresponding estimator under the restriction implied by c for the expected value of the treatment condition will be denoted by:*

$$y\widehat{(\mathcal{S}, \mathcal{S}')}_w(c)$$

The associated quantity of interest is:

$$\tau\widehat{(\mathcal{S},\mathcal{S}')}_w(c) = y\widehat{(\mathcal{S},\mathcal{S}')}_w(c) - y\widehat{(\mathcal{S}',\mathcal{S})}_w(c)$$

The analysis requires one important assumption, what will be referred to as the *non-cascading assumption.* Intuitively, this means that: a) the distance measure is well-behaved–more restrictive assumptions on distance imply that the observed outcome is closer to the desired potential outcome; and b) it cannot be the case that assigning any unit i to $d_i \neq 0$ implies that the potential outcome for every unit is affected, i.e. there are no spillovers between units that are sufficiently far from each other.[11]

Theorem 3.4.7. *Let*

$$\max_{\mathbf{d}' \in \tilde{Q}_c^i(\mathbf{d})} |y_i(\mathbf{d}') - y_i(\mathbf{d})| = \alpha_{i,\mathbf{d}}(c)$$

Then:

$$|E(\widehat{y(\mathcal{S},\mathcal{S}'_i)_w})(c) - \overline{y(\mathcal{S},\mathcal{S}'_i)_w}| < \frac{\sum\limits_{i\in\Omega} w_i \sum\limits_{\mathbf{d}\in\mathbf{S}_i} \sum\limits_{\mathbf{d}'\in\mathbf{S}'_i} \alpha_{i,\mathbf{d}}(c)}{\sum\limits_{i\in\Omega} |\mathbf{S}_i||\mathbf{S}'_i| w_i}$$

Assumption 3.4.8 (Non-Cascading Assumption)**.** *Let*

$$\max_{\mathbf{d}' \in \tilde{Q}_c^i(\mathbf{d})} |y_i(\mathbf{d}') - y_i(\mathbf{d})| = \alpha_{i,\mathbf{d}}(c)$$

As $c \to \infty$, $\alpha_{i,\mathbf{d}}(c) \to 0$ for all i and $\mathbf{d} \in \mathbf{S}_i \cup \mathbf{S}'_i$.

Corollary 3.4.9 (Convergence in Expectation)**.** *Under assumption 3.4.8, $E(\widehat{y(\mathcal{S},\mathcal{S}'_i)_w}(c)) \to \overline{y(\mathcal{S},\mathcal{S}'_i)_w}$ as $c \to \infty$. Define a sequence $(c_k)_{k=1}^{\infty} \to \infty$. It follows that for every $\varepsilon > 0$, there exists K such that $|E(\widehat{y(\mathcal{S},\mathcal{S}'_i)_w}(c_K)) - E(\widehat{y(\mathcal{S},\mathcal{S}'_i)_w}(c_k))| < \varepsilon$ for all $k > K$.*

[11] This idea may be somewhat testable if one considers the "uniformity trial" framework discussed by Rosenbaum (2007). One can use either a pre-measurement or a placebo trial to estimate the potential outcome under the no treatment condition, and compare the potential outcomes of untreated units in the actual experiment.

Although the proof is relatively straightforward, it has deep implications for the analytic framework. In particular, it shows that particular units i that may be highly influenced or are highly influential in terms of spillovers are unlikely to make a large impact on the bias. Thus, while the non-cascading assumption is a pointwise convergence property, the estimator behaves much like uniform convergence when there are few influential or highly influenced units. This also suggests some basic guidelines in selecting a social distance for the analysis. The goal is to select a distance measure for which those units that interfere the most with observing the potential outcome of i are seen as "close" to i. While mathematically any social distance measure that satisfies the non-cascading assumption is admissible, it will typically be necessary to find a social distance that quickly "removes" the bias due to spillovers due to a strong negative correlation to the magnitude of spillovers.

Non-Interference and Social Distance in the Example

The toy example affords an opportunity to clarify some of the complexities of the above exposition. In particular, consider a simulated experiment where exactly 10 of the 50 experimental units, in the same configuration shown in figure **??**, are treated (with equal probability). A natural choice for a non-interference assumption over a social network is the network distance, the minimum number of edges between two nodes. Formally, the minimum path length between two nodes, i and j, defines the distance measure, $\rho(i, j)$. If $c = 1$, the potential outcomes of a unit may vary with the treatment status of its neighbors, but not with the treatment status of the neighbors of its neighbors. If $c = 2$, a neighbor of a neighbor may interfere with the potential outcome of a unit, and so on. The classic non-interference assumption has $c = 0$. Of course, if the nodes were defined with other features, social distances (continuous or discrete) could be defined with respect to those features (e.g., geographic distance, frequency of interaction, friendship). This network distance formulation is formalized below.

Definition 3.4.10. *A **path** is an alternating sequence of edges and nodes where each node is incident with the edge before and after it in the sequence. The **shortest path** between two* *nodes i and j, $\rho(i,j)$, is the path connecting i and j with smallest number of edges.*

Definition 3.4.11. *The restriction c, is a number such that two nodes, i and j, with $\rho(i,j) \leq c$, implies:*

$$\mathbf{d}' \notin \tilde{Q}^i_c(\mathbf{d}) \quad \textit{if } d_j \neq d'_j$$

The resulting non-interference assumption induced by c corresponding to a reference assignment, $\mathbf{d}^*$, is:

$$\tilde{Q}^i_c(\mathbf{d}^*) = \{\mathbf{d}|d_k = d^*_k \textit{ if } \rho(i,k) \leq c\}$$

Using this definition, the non-interference assumptions for each value c are:

$$\tilde{Q}^i_c(\mathbf{d}_j) = \{\mathbf{d}|d_j = 1,\ d_k = 0 \text{ if } j \neq k \text{ and } \rho(i,k) \leq c\} \tag{3.4.4}$$

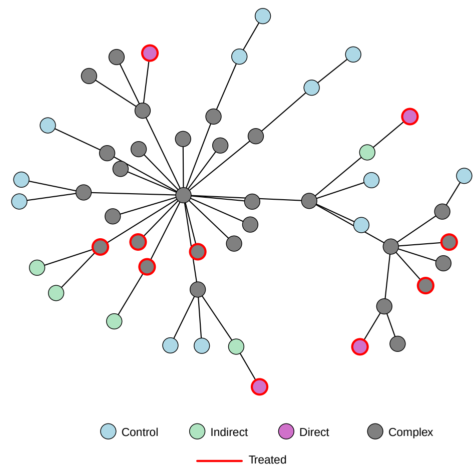
$$\tilde{Q}^i_c(\mathbf{0}) = \{\mathbf{d}|d_k = 0 \text{ if } \rho(i,k) \leq c\}$$

Figure 3.3 depicts how the choice of c affects observed potential outcomes. In figure 3.3, 10 of 50 units are selected to receive treatment (those units with a thick red border), while no other units receive treatment. The magenta units correspond to units for which $y_i(1)$ observed, and the light blue units correspond to units where $y_i(0)$ is observed. Finally, $y_{ij}(0,1)$ is observed among the light green units, those units with a single treated neighbor. Finally, complex outcomes are represented by gray units, which do not correspond to any of the outcomes of interest. Note that $y_{ij}(1,1)$ is never observed in these graphs.

Increasing $c = 1$ to $c = 2$ makes it more difficult to observe the outcomes of interest. Only two untreated units cannot be classified as an outcome of interest when $c = 1$, but 23 unclassified outcomes for untreated units and 6 unclassified outcomes for treated units emerge when c is increased to 2. For example, six treated units cannot be classified when $c = 1$ is increased to $c = 2$ because they are separated by a path length of 2. In general, stipulating more stringent non-interference assumptions through a higher c, or increased network density will yield more units observed with complex spillovers. Thus, under complete

Figure 3.3: Observing Potential Outcomes with $c = 1$ and $c = 2$

(a) $c = 1$



(b) $c = 2$

Figure 3.3 describes how selecting $c = 1$ and $c = 2$ affects the ability to observe various potential outcomes, with 3.3(a) depicting $c = 1$, and 3.3(b) depicting $c = 2$. In each subfigure, a graph with 50 units, where 10 have been selected for treatment, is depicted. Those units with a thick red border correspond to units selected for treatment, while the rest of the units are not selected for treatment. The magenta units are those units for which $y_i(1)$ is cleanly observed, and the light blue units are those for which $y_i(0)$ is cleanly observed. The light green units are those for which $y_{ij}(0, 1)$ is cleanly observed, i.e., the spillover to an untreated unit , i from a single treated neighbor, j, is observed without interference. $y_{ij}(1, 1)$ is never observed in the graphs. The gray units correspond to complex spillovers where none of the potential outcomes of interest are observed.

random assignment, estimation is likely to get increasingly inefficient as c or network density grows.

3.4.3 The Relationship Between Distance Restrictions and the Admissible Sample

In this subsection, the toy example discussed above is extended to a more realistic setting. In particular, a network containing 10,000 nodes is constructed from the Barabási-Albert algorithm with $m = 2$. Furthermore, 250 nodes are randomly selected to receive the treatment, and all the quantities of interest are constructed as above. The first set of considerations involves evaluating the tradeoff between the admissible sample and the distance restrictions placed on the calculation. The admissible sample is defined as the set of nodes that may be placed in each treatment condition of interest with positive probability under the distance restriction under consideration. In order to generate these probabilities, 250 nodes were selected for treatment under complete randomization over 5,000 draws and the respective treatment conditions (with respect to the distance restriction, c) were calculated in each draw. The probability of assignment to any treatment condition for a node, conditional on the distance restriction, is simply the fraction of draws over which that node is observed to be in the treatment condition. Table 3.1 summarizes the results for the admissible sample.

	Admissible Sample
$c = 1$	9983
$c = 2$	5838
$c = 3$	394

Table 3.1: Admissible Sample

Table 3.1 demonstrates a strong tradeoff between the admissible sample and the distance restriction, c. Once c is increased to 3, only 4% of the original nodes are included in the admissible sample. There are a number of ways of boosting the sample when more stringent restrictions are considered. First, the researcher may choose to estimate different admissible samples for different quantities of interest, which would boost the sample since a node is not required to be observed in each treatment condition of interest. The downside of this approach is that the units are no longer comparable across quantities of interest. A second, more complicated approach is to incorporate the distance restrictions explicitly in the sampling algorithm to guarantee that more units will be observed in each treatment condition by minimizing the number of units experiencing complex spillovers at a particular distance restriction. This is a significantly more complicated task than the complete randomization considered above, but it yields strong gains in the efficiency of the estimation procedure. This approach, as well as other design phase criteria, are discussed at length in Coppock and Sircar (2014).

Nonetheless, the researcher must balance generally conservatism in terms of the distance restriction and the representativeness of the sample. Figure 3.4 shows the ratio of the degree distribution[12] under $c = 2$ to the degree distribution under $c = 1$ (for degrees ≤ 21). The ratio shrinks at higher degrees, suggesting that higher distance restrictions underrepresent nodes with higher degrees. Under $c = 2$, only one node of degree greater than 31 is included in the admissible sample (not shown), whereas there are 78 such nodes in the original sample. In order to address this issue, the researcher may choose to define reference assignments where a fixed set of higher degree nodes are guaranteed to be treated/untreated and the complete randomization is conducted over lower degree nodes. Alternatively, the researcher may define treatment conditions with less restrictive assumptions for higher degree nodes (e.g., the node is treated/untreated without regard for the number of treated neighbors).

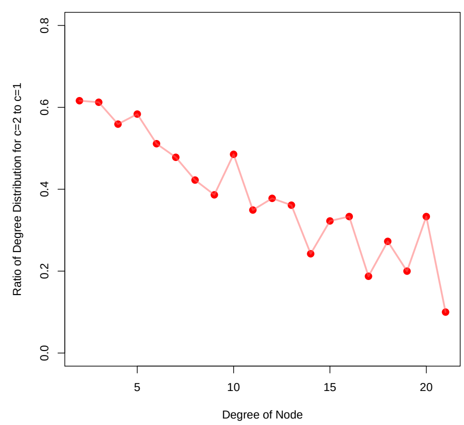


Figure 3.4: Ratio of Degree Distribution under $c = 2$ to $c = 1$

Figure 3.4 shows the ratio of the degree distribution under $c = 2$ to the degree distribution under $c = 1$.

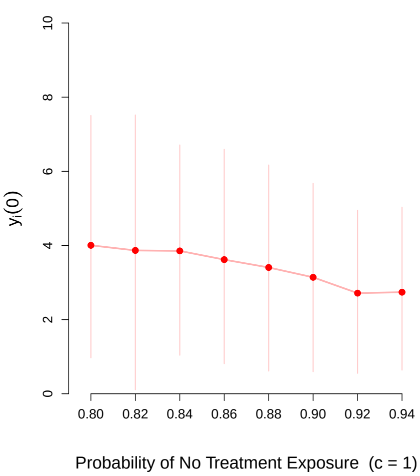
In the previous section, the framework was defined with regard to the probability of assignment to each treatment condition. At first, this might seem odd since the researcher is unlikely to randomize with respect to any characteristics of the nodes. However, the very structure of the network affects the probability of observing each treatment condition. In order to see this, figure 3.5 plots the values of the potential outcomes[13] associated with the

[12]The degree distribution is defined as the distribution of the degrees of the nodes in a network.

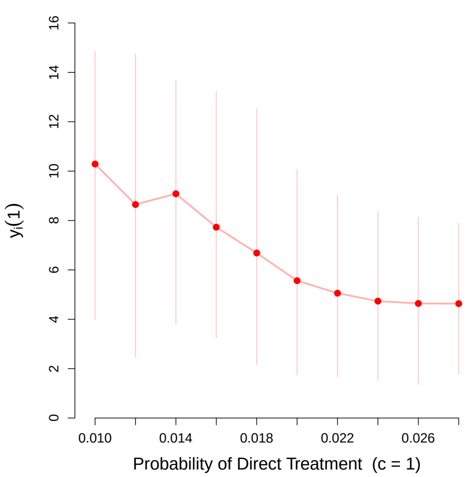
[13]The potential outcome is drawn as a function of the degree of the node. More information is provided in

direct treatment condition, $\mathcal{S}_1$, and untreated condition, $\mathcal{S}_0$, against their respective probabilities of assignment to the treatment condition under $c = 1$. A clear relationship is observed in both cases, with directed treatment condition demonstrating a sharper relationship due to the fact that both untreated potential outcome and treatment effect for a node are drawn as a function of its degree. Taken together, figure 3.5 demonstrates the need for addressing the probabilities of assignment to treatment conditions for estimation of causal effects under spillovers. The researcher is recommended to plot the probabilities of assignment to a treatment condition against observed outcomes in the treatment condition to understand this relationship.

Figure 3.5: Correlation between Probability of Assignment and Potential Outcome



(a) Direct Treatment Condition



(b) Untreated Condition

Figure 3.5 describe how the potential outcomes in the directly treated [3.5(a)] and untreated [3.5(b)] conditions vary as a function of their respective probabilities of assignment to the treatment condition. The lighter colored bands represent 90% confidence intervals within a $\pm$ .001 ($\pm$ .01) band for the directly treated (untreated) condition, and point estimate represents the mean over the band.

3.4.4 Generality of the Social Distance Approach

In this paper, the goal is to draw causal inferences about quantities of interest without making parametric or functional assumptions about potential outcomes in the presence of spillovers.

the next section.

Although the social distance approach is more general than approaches that assume known structures for spillovers, one might wonder about the existence of more general approaches to causal identification in this setting.

It is possible to show that the social distance approach characterizes a fully general framework. To see why, first consider a situation where the analyst knows that spillovers from each treated unit reaches every other unit in the population. In this scenario, the analyst cannot directly observe more than one treatment condition of interest, whereas the quantity of interest requires two such conditions to be observed. More concretely, consider some unit $u \in \Omega$. Since spillovers emanating from u reach every other unit, the outcome observed for some other unit v is always dependent upon the treatment status of u; that is, for $v \in \Omega \setminus \{u\}$, $y_v|d_u = a_u \neq y_v|d_u = a'_u$ when $a_u \neq a'_u$. Since this is true for each unit $u \in \Omega$, each constellation of treated units admits a unique potential outcome for each unit in the population.

Causal identification in this setting, thus, requires units for which spillovers will not reach every other unit in the population, as well as units for which there exists some unit from spillovers do not reach. Consider a function ρ and let $u, v \in \Omega$. Define $\rho(u, u) = 0$, $\rho(u, v) = 1$ if spillovers from v reach u and $\rho(u, v) = 2$ if spillovers from v do not reach u. It is easy to verify that ρ is a quasi-metric that satisfies the non-cascading assumption. Therefore, at a minimum, the analyst requires assumptions that are mapped to a social distance metric satisfying the non-cascading assumption.

In a non-parametric setting, the maximum amount of information an analyst can bring to the problem is the relative magnitude [as opposed to absolute magnitude] of spillovers between units. Now consider a function $\rho^* : \Omega \times \Omega \to \mathbb{R}$ such that for each $u, u', v, v' \in \Omega$, $\rho^*(u, v) \leq \rho^*(u', v')$ implies that the magnitude of the spillover from v to u is greater than or equal to of that from v' to u'. Furthermore, for all $u, v \in \Omega$, $u \neq v$, define c^* such $\rho^* > c^*$ implies that spillovers do not exist between the units under consideration. The non-cascading assumption requires that for each $u \in \Omega$, there exists some $v, v' \in \Omega$, where $\rho^*(u, v) > c^*$ and $\rho^*(v', u) > c^*$. Now let $\min \rho^* \geq \underline{c}$ and $\max \rho^* \leq 2\underline{c}$; this can always be done through simple transformations that preserve the order for some candidate social distance function that meet the above conditions. It can be easily verified that ρ^* is a quasi-metric that satisfies the non-cascading assumption. This shows that any information that be can brought to bear on the problem of spillovers in a non-parametric setting can be stated in terms of a social distance metric. A "true" social distance metric will be any quasi-metric $\tilde{\rho}$ such that for each $u, u', v, v' \in \Omega$, $\rho^*(u, v) \leq \rho^*(u', v')$ implies $\tilde{\rho}(u, v) \leq \tilde{\rho}(u', v')$.

However, the researcher will rarely possess the necessary information to determine a true social distance metric, and not all admissible social distance measures will always be usable or available. For instance, the social distance metric may only be weakly correlated to the bias, so it is relatively ineffective in removing such bias. Furthermore, the actual sample may exhibit small deviations from the non-cascading assumption. In order to address these concerns, the following section develops more stringent assumptions on the hypothesized social distance measure (that the expected bias as a function of the social distance measure is decreasing and convex) in the estimation strategy. These more stringent assumptions allow the researcher to generate a bound on the bias in estimation.

3.5 Analytical Framework

This section demonstrates how successive distance restrictions can be used in a practicable empirical framework to estimate causal quantities of interest in the presence of spillovers. In particular, an example of the estimation strategy is shown, as well as an intuitive test to deduce convergence of the estimator with respect to distance restrictions. Given the complex pattern of probabilities of assignment to various treatment conditions, and the challenges of efficiency and power in the spillover setting, this section proposes the use of cubic thin-plate regression splines (TPRS) to flexibly estimate quantities of interest. The TPRS are estimated using Markov Chain Monte-Carlo (MCMC), so statistical inference results from a straightforward application of the Bayesian inferential machinery. The argument in this section follows two steps. First, more stringent, but intuitive, assumptions are placed on the social distance measures in order to create a bound on the bias in estimation, and it is shown that the preferred "measure" is that of the expected sample size at each restriction. Second, the TPRS estimator, and resulting inferential framework, is developed over this social distance measure.

The arguments in this section may be viewed as supporting the so-called "calibrated Bayes" approach (Little, 2006, 2011), which seeks to develop flexible regression models with the benefits of Bayesian estimation and desirable frequentist properties. However, unlike traditional regression approaches where researchers can calibrate the "fit" of the model to the data, no such fit may be analyzed with respect to the theoretical parameters of interest in this paper. For this reason, the strategy must incorporate flexible estimation over increasingly stringent distance restrictions and develop a test to deduce "convergence."

Inference in this Bayesian framework is straightforward since it only requires a comparison of the posteriors generated from the model and is based upon credible intervals. By contrast, it is somewhat unclear how to proceed in a frequentist framework and create classical confidence intervals. Fisher-style randomization inference is hampered by several treatment conditions of interest and the fact that for many units no treatment condition of interest will be observed. Furthermore, as Aronow and Samii (2013) show, design-consistent estimation of the variance of standard inverse-probability weighted (IPW) estimators, the usual estimator in experimental settings, is typically not possible.

Section 5.1 develops the form of the social distance measure that will be used in real empirical settings. Section 5.2 develops the proposed TPRS estimator, section 5.3 discusses the simulated data, and section 5.4 describes the inferential framework with sample calculations. Finally, section 5.5 compares the TPRS estimator to more classic inverse-probably weighted (IPW) estimators, typically used in experimental settings, through simulation. It is shown that TPRS provides superior performance in terms of root-mean squared error (RMSE), with a reduction of approximately 30% in RMSE.

3.5.1 Developing a Plausible Social Distance Measure

As discussed in the previous section, defining a usable social distance measure for empirical analysis can be difficult. In order create a practicable measure, it will be important to address two issues.

First, while only the non-cascading assumption is required for appropriate estimators to converge to the quantity of interest, without more stringent assumptions the expected bias may exhibit serious non-monotonicities (and non-convexities) with respect to the social distance measure. In other words, for the purposes of analysis, it is desirable to choose a social distance measure such that the function mapping the social distance measure to expected bias is strictly decreasing and convex.

Second, it is important to select a social distance measure that is sensitive to the sample size. In other words, it is entirely plausible that a more stringent distance restriction barely affects the sample under consideration. In order to explicitly tether the sample size to the social distance measure, the social distance measure will be taken to be the expected sample size under each distance restriction.

Assumptions on Social Distance Measure

The following assumption of a decreasing, convex function relating expected bias and the social distance measure will be necessary to make calculations.[14]

Assumption 3.5.1 (Decreasing, Convex Bias Function)**.** *Consider an unbounded increasing sequence of distance restrictions,* $(c_k)_{k=0}^{\infty}$*. Define the expected bias function:*

$$B(c_k) = |E(\widehat{y(\mathcal{S}, \mathcal{S}'_i)_w})(c_k) - \overline{y(\mathcal{S}, \mathcal{S}'_i)_w}|$$

This expected bias function is decreasing and convex. In particular:

1. $B(c_k) \leq B(c_j)$ *if* $j < k$

[14] I am particularly grateful to Macartan Humphreys for pointing out importance of these assumptions for this framework.

2. *For every* $0 \leq \lambda \leq 1$, $B(\lambda c_k + (1-\lambda)c_j)) \leq \lambda B(c_k) + (1-\lambda)B(c_j)$ *for* $j \neq k$

Although the following assumption is not necessary, it is natural to assume that the bias is always going in the same direction. The following assumption states this formally:

Assumption 3.5.2 (Same Direction Bias)**.** *The bias of the estimator will always be in the same direction. That is,*

$$|E(\widehat{y(\mathcal{S}, \mathcal{S}_i')_w})(c_j) - E(\widehat{y(\mathcal{S}, \mathcal{S}_i')_w})(c_k)| \leq B(c_j); \quad k > j$$

Note that the non-cascading assumption implies that $B(c_k) \to 0$ as $k \to \infty$. Using these definitions and assumptions, one can put a bound on the bias at the distance restriction c_K, given an assumption about the value of $B(c_k)$. In particular, the decreasing convexity assumptions imply that for $c_j < c_K < c_k$, the following holds:

$$B(c_K) - B(c_k) \leq \frac{c_k - c_K}{c_K - c_j}(B(c_j) - B(c_K)) \tag{3.5.1}$$

This follows from setting $\lambda = \frac{c_K - c_j}{c_k - c_j}$ and solving through the inequality, $B(\lambda c_k + (1 - \lambda)c_j) \leq \lambda B(c_k) + (1-\lambda)B(c_j)$.

The next theorem shows that this inequality implies that convergence of the estimator for the expected value of the treatment condition over distance restrictions implies an unbiased estimate. It is for this reason that it will be desirable to demonstrate convergence of the estimator.

Theorem 3.5.3 (Duality Between Convergence and Unbiasedness)**.** *Consider the expectation of an unbiased estimator,* $E(\widehat{y(\mathcal{S}, \mathcal{S}_i')_w})(.)$, *and assume the assumptions in 3.5.1 and 3.5.2. If the estimator converges; that is, for an unbounded increasing sequence of distance restrictions,* $(c_k)_{k=0}^{\infty}$*:*

$$E(\widehat{y(\mathcal{S}, \mathcal{S}_i')_w})(c_j) = E(\widehat{y(\mathcal{S}, \mathcal{S}_i')_w})(c_K) \quad \textit{for some } K > j$$

then

1. $E(\widehat{y(\mathcal{S},\mathcal{S}'_i)_w})(c_j) = E(\widehat{y(\mathcal{S},\mathcal{S}'_i)_w})(c_k)$ *for all* $k \geq K$

2. $B(c_k) = 0$ *for all* $k \geq K$

Proof: Note that:

$$E(\widehat{y(\mathcal{S},\mathcal{S}'_i)_w})(c_k) = y(\mathcal{S},\mathcal{S}'_i)_w + B(c_k) \text{ or } E(\widehat{y(\mathcal{S},\mathcal{S}'_i)_w})(c_k) = \overline{y(\mathcal{S},\mathcal{S}'_i)_w} - B(c_k); \quad B(c_k) \geq 0 \tag{3.5.2}$$

If $E(\widehat{y(\mathcal{S},\mathcal{S}'_i)_w})(c_j) = E(\widehat{y(\mathcal{S},\mathcal{S}'_i)_w})(c_K)$, then $|E(\widehat{y(\mathcal{S},\mathcal{S}'_i)_w})(c_j) - E(\widehat{y(\mathcal{S},\mathcal{S}'_i)_w})(c_K)| = 0$ Plugging in the equations in (3.5.2), it follows that $|B(c_j) - B(c_K)| = 0$.

If $B(c_j) - B(c_K) = 0$, then consider (3.5.1). It follows that $B(c_j) - B(c_k) = 0$ for $k \geq K$. It also follows, from l'Hôpital's Rule, $\lim_{k\to\infty} B(c_j) - B(c_k) = B(c_j) - \lim_{k\to\infty} B(c_k) = 0$. The non-cascading assumption implies $\lim_{k\to\infty} B(c_k) = 0$, so it follows that $B(c_j) = 0$ and that $B(c_k) = 0$ for $k \geq K$. Plugging this back into (3.5.2) proves the theorem. □

One can also use (3.5.1) to generate a maximum bound on the bias in estimation based on the statistical power of the procedure from assumptions 5.1 and 5.2. In particular, let ΔB_{jk} be the *minimum detectable difference in bias*. That is, given restrictions c_j and c_K, ascertain the minimum value of $B(c_j) - B(c_K)$ that can be detected from the statistical procedure. Then, plugging into (3.5.1) yields:

$$B(c_K) \leq \frac{c_k - c_K}{c_K - c_j}\Delta B_{jk} + B(c_k) \tag{3.5.3}$$

While (3.5.3) is just a simple rearranging of terms, it can yield important insights in a power analysis and do much to convince the reader of the quality of the estimate. The bound requires some assumption about the magnitude of the bias at $B(c_k)$, but at a relatively stringent distance restriction one may reasonably assume the value to be near zero (or one may perform sensitivity analysis on the values for c_k and $B(c_k)$).[15] Under this assumption, and in an experimental setting with sufficient to guarantee that ΔB_{jk} is relatively small, the magnitude of the bias at $B(c_K)$ is likely to be small.

Tethering the Social Distance Measure to the Sample Size

One way in which the researcher may "cheat" is to consider more restrictive distance restrictions that do not dramatically alter the sample under consideration. For instance, for certain treatment assignment vectors in the example, there may be little difference in the sample under consideration for $c = 0$ and $c = 1$. In such a situation, it may seem that estimator has converged, when in reality the samples under consideration are very similar. Such a scenario is also likely to cause non-convexities in the distance measure.

In order to prevent this situation, one can tether the social distance measure to the sample sizes under consideration. More precisely, let $E(N(\overline{y(\mathcal{S}, \mathcal{S}')}, c_k))$ be the expected sample size under consideration when the sample is restricted to a specific distance restriction for a particular treatment condition of interest. Define an induced distance restriction:

$$c'_k = E(N(\overline{y(\mathcal{S}, \mathcal{S}')}, 0)) - E(N(\overline{y(\mathcal{S}, \mathcal{S}')}, c_k))$$

These new social distance measures are a form of "convexification" of the existing social distance measure because small/no changes in the sample size for significant changes in the existing social distance measure may create non-convexities. In these new sample size based social distance measures, the greatest restrictions correspond to those situations where the existing social distance measure induces very few units to be observed in the treatment condition of interest. If the researcher believes that there is still a significant bias present under such stringent distance restrictions, then this implies that the sample is not close to satisfying the non-cascading assumption.

[15] As will be shown below, such an assumption will have intuitive meaning when the distance measure is tethered to the sample size.

3.5.2 Developing the Estimator

This sections develops the TPRS estimator for experimental analysis under spillovers. The process entails iteratively fitting TPRS and checking for convergence at the end of the process. In particular, a TPRS is fit to the data with respect to the probability of assignment induced by the weakest restriction and the residuals are obtained. In the next iteration, the residuals corresponding to those observations satisfying a stronger distance restriction are taken as the dependent variable, and TPRS is fit to these residuals with respect to the probabilities of assignment induced by this stronger distance restriction. The researcher continues this process until convergence, where convergence is achieved when the TPRS approximates a zero function.

TPRS is a penalized spline technique which fits naturally into a Bayesian regression framework and can be easily estimated through open source software without serious difficulty. While other flexible semi-parametric or data mining methods may also work in this context, the main benefits of this approach are its simplicity, access to a natural inferential framework and ability to prevent serious overfitting of the data. TPRS was first introduced by Wood (2003) and extended to a fully Bayesian implementation (in open source software) by Crainiceanu et al. (2005b), which the exposition below follows closely.

Let $\pi_i(c)$ denote the probability of assignment to a treatment condition under the distance restriction c, and let y_i be the corresponding observed outcome. The goal is to determine the "smooth" function m such that satisfies:

$$y_i = m(\pi_i(c)) + \varepsilon_i; \quad \varepsilon_i \sim N(0, \sigma_\varepsilon^2) \tag{3.5.4}$$

Under this smoothness assumption, the weighted average expected value of condition $\mathcal{S}$ with respect to $\mathcal{S}'$ induced by the restriction c is given by:

$$\widehat{y(\mathcal{S}, \mathcal{S}')}_w(c) = \frac{1}{\sum_{i \in \Omega} |\mathbf{S}_i||\mathbf{S}'_i| w_i} \sum_{i \in \Omega} \sum_{\mathbf{d} \in \mathbf{S}_i} |\mathbf{S}'_i| w_i m(\pi_i(c)) \tag{3.5.5}$$

In order to estimate the function m, TPRS is used. TPRS addresses several issues which are beyond the scope of this paper, including admitting a low-rank spline which has low sensitivity to number and placement of knots. Consider a sequence of knots $\kappa_1, \ldots, \kappa_K$. The cubic TPRS to the function m, $\widehat{m}$, can be written as:

$$m(\widehat{\pi_i(c), \boldsymbol{\theta}}) = b_0 + b_1\pi_i(c) + \sum_{k=1}^{K} g_k|\pi_i(c) - \kappa_k|^3 \tag{3.5.6}$$

It will be easier to rewrite the equation in matrix form:

$$m(\widehat{\pi_i(c), \boldsymbol{\theta}}) = \mathbf{X}\mathbf{b} \ + \ \mathbf{Z}_K\mathbf{g} \tag{3.5.7}$$

where $\boldsymbol{\theta} = \{b_0, b_1, g_1, \ldots, g_K\ \}$, $\mathbf{X}_i = (1, \pi_i(c))$, and $\mathbf{Z}_{Ki} = (|x_i - \kappa_1|^3, \ldots, |x_i - \kappa_K|^3)$. For the "smoothing parameter," λ, the penalized approach yields the parameter estimates $\widehat{\boldsymbol{\theta}}$:

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \ \sum_{i=1}^{N}(y_i - m(\pi_i(c)))^2 + \frac{1}{\lambda}\boldsymbol{\theta}^{\boldsymbol{T}}\mathbf{G}\boldsymbol{\theta} \tag{3.5.8}$$

where $\mathbf{G}$ is a positive-definite penalty matrix defined by:

$$\mathbf{G} = \begin{bmatrix} \mathbf{0}\ \mathbf{0} \\ \mathbf{0}\ \boldsymbol{\Gamma} \end{bmatrix}$$

with $\boldsymbol{\Gamma}$ a $K \times K$ matrix with the (j,k) entry defined as $|\kappa_j - \kappa_k|^3$.

Now, let $\mathbf{Z} = \mathbf{Z}_K\boldsymbol{\Gamma}^{-\frac{1}{2}}$ and $\boldsymbol{\gamma} = \boldsymbol{\Gamma}^{\frac{1}{2}}\mathbf{g}$

Ruppert et al. (2003) and Crainiceanu et al. (2005b) have shown the equivalence of (3.5.8) with parameters resulting from estimation of the following Bayesian mixed model:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \tag{3.5.9}$$

$$\gamma_i \sim N(0, \sigma^2_\gamma) \quad \varepsilon_i \sim N(0, \sigma^2_\varepsilon)$$

3.5.3 Data

The method is demonstrated on simulated data. The simulation is generated to be realistic but also to maximize difficulties that might arise during estimation. Since the probability of assignment to each condition is likely to be a function of the network, potential outcomes at the level of the node are drawn as a function of node's degree. Furthermore, the extent of complex spillovers are constructed to be a function of relative "centrality" of the node, where each node receives spillovers from nodes that are within a network distance of 4, and the magnitude of spillovers is constructed be extremely large by most empirical standards. As before, a fixed scale-free network with 10,000 nodes is drawn according the Barabási-Albert algorithm with $m = 2$, which guarantees that the minimum degree for a node is 2 and that the network is fully connected.

The potential outcomes for the untreated condition are generated as a function of degree, δ_i:

$$y_i(\mathbf{0}) \quad \sim \quad N\left(2 + \log(\delta_i),\ \frac{1}{4}(2 + \log(\delta_i))^2\right)$$

The direct treatment effect for each unit i is also generated as a function of δ_i and denoted by $\tau_i(\mathbf{d}_i)$:

$$\tau_i(\mathbf{d}_i) \quad \sim \quad N\left(2 + \log(\delta_i),\ \frac{1}{9}(2 + \log(\delta_i))^2\right)$$

$$y_i(\mathbf{d}_i) = y_i(\mathbf{0}) + \tau_i(\mathbf{d}_i)$$

Spillovers were modeled as a function of the network adjacency matrix, A, which is a symmetric matrix where the entry in row i and column j, $(A)_{ij}$, takes the value 1 if there exists an edge between nodes i and j, and 0 otherwise. For A^K, it follows that the entry in row i and column j, $(A^K)_{ij}$, denotes the number of paths of length K between nodes i and j. If treated, the direct treatment effect of node j "spills over" to node i according to the function $\kappa_i * \phi_j^K \tau_i(\mathbf{d}_j) * (A^K)_{ji}$.

For a given assignment vector $\mathbf{d}^*$, spillovers were generated if $K \leq 4$ with $\phi_j = 0.2$ and $\phi_j = 0.6$ and $\kappa_i = 1$ if $\mathbf{d}_i^* = 1$. This captures the following realistic conditions: 1) spillovers move a long way (and thus the estimate under the most restrictive condition $c = 2$ is not unbiased); and 2) a significant portion of the direct effect of the treatment spills over the network. The complexity of this process provides a good test of the proposed estimator. The observed outcome for unit i is given by:

$$y_i^{obs}(\mathbf{d}^*) = \mathbf{d}_i^* * y_i(\mathbf{d}_i) + (1 - \mathbf{d}_i^*) * y_i(\mathbf{0}) + \sum_{K=1}^{4} \sum_{j \in \Omega} \phi_j^K * \mathbf{d}_j^* * \tau_i(\mathbf{d}_j) * (A^K)_{ji} \qquad (3.5.10)$$

A relatively small number of units, 250, out of a total of 10,000 units were treated to minimize the extent of complex spillovers in the data and allows for more frequent observation of treatment conditions of interest. Although this paper has not discussed optimal designs in a spillover setting, one might reasonably expect the researcher to take a similar approach and directly treat a fairly small percentage of a large number number of units. In the situation where the researcher chose less optimal designs than the one presented here, one might reasonably expect even greater efficiency benefits from the TPRS approach due to smaller samples sizes in treatment conditions of interest. Tests are currently being conducted in a higher bias environment, where a higher percentage of units are selected for assignment to treatment.

3.5.4 The Inferential Procedure

The probability of assignment to each treatment condition under each distance restriction was calculated from 5000 simulations [complete random assignment of 250 nodes out of 10,000 nodes] in a manner described in the previous section. In this section, the method is demonstrated with respect to estimating the direct treatment condition and the indirect exposure condition. Once the probability of assignment under each distance restriction is obtained, TPRS can be fit to the data. This subsection describes statistical inference resulting from this framework.

Sample Calculation for Direct Treatment Condition

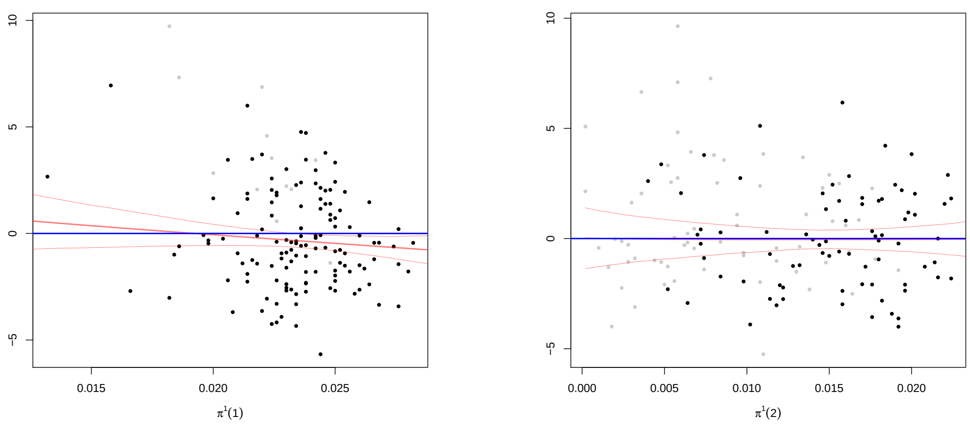
In this analysis, as in the previous section, the social distance considered is network distance. In order to estimate the effect, and assess convergence, the procedure is considered at distance restrictions $c = 0, 1, 2$ to estimate the average of the direct treatment condition. The data are generated with $\phi_j = 0.6$, which means 60% of a unit's treatment effect spills over to neighbors.

Let $\mathbf{y}^1(c)$ denote the vector of directly treated units observed under restriction c with individual observations $y_i^1(c)$, and let $\widehat{m_c}$ denote the estimated mean function of $y_i^1(c)$ resulting from the TPRS fit with respect to assignment probability to direct treatment, $\pi_i^1(c)$. The estimation strategy follows the following steps, starting with $c = 0$, and fitting iteratively up to $c = 2$. (or until convergence).

1. Calculate mean function, $\widehat{m_c}$, of $y^1(c)$
2. Form the residual $\varepsilon_i^1(c+1) = y_i^1(c+1) - m_c\widehat{(\pi_i^1(c)})$, the residual for treated units that satisfy the restriction $c+1$ with respect to the mean function corresponding to c.
3. Fit the TPRS with dependent variable $\varepsilon_i^1(c+1)$ and predictor $\pi_i^1(c+1)$, yielding function $\widehat{\eta_{c+1}}$
4. Form $\widehat{m_{c+1}} = \widehat{m_c} + \widehat{\eta_{c+1}}$
5. Test for convergence by checking if $y_i^1(c+1) - m_{c+1}\widehat{(\pi_i^1(c+1)})$ approximates a zero function
6. Repeat steps 2-5 until convergence is reached

One example of the calculation is conducted here. The true value of the average of the direct treatment condition is 4.95. At $c = 0$, the estimate of average of the direct treatment condition, $\widehat{m_0}$ is the naive mean of treated units, which is 5.28 in this case. For $c = 1$, the residual $\varepsilon_i^1(1)$ is constructed by taking the observed values of $y^1(1)$ and subtracting off $\widehat{m_0}$. Now, TPRS is fit to $\varepsilon_i^1(1)$ with the predictor $\pi_i^1(1)$. The fitted spline, $\widehat{\eta}_1$, with 90% posterior interval, is shown in figure 3.6(a).

Figure 3.6: Fitted Splines for Direct Treatment Condition



(a) Fit to Residual under $c = 1$

(b) Fit to Residual under $c = 2$

It is evident that the spline is downward sloping, and a significant portion of the 90% posterior interval fails to intersect with the zero function, suggesting that the estimator has not reached convergence. Now, $\widehat{m_1} = \widehat{m_0} + \widehat{\eta_1}$ is constructed and the estimated expectation of the treated condition under $c = 1$ is estimated from the predicted values as:

$$\widehat{y(\mathcal{S}_1, \mathcal{S}_0)}(1) = \frac{1}{N} \sum_{i \in \Omega} \widehat{m_1(\pi_i^1(1))} = 4.97$$

Since convergence has not been reached, the restriction $c = 2$ is considered. Again, a residual under the restriction of $c = 2$ is formed as $\varepsilon_i^1(2) = y_i^1(2) - \widehat{m_1(\pi_i^1(1))}$, and the TPRS is fit with the predictor $\pi_i^1(2)$. The resulting spline, η_2 with 90% posterior interval is shown in figure 3.6(b). This spline seems very close to the zero function, suggesting convergence. Since this suggests convergence, we take the value $\widehat{y(\mathcal{S}_1, \mathcal{S}_0)}(2)$ as the estimate of the average of the direct treatment condition. Once again, the estimate is calculated through fitted values. First, the mean function is constructed, $\widehat{m_2} = \widehat{m_1} + \widehat{\eta_2}$. Then, the the estimate of the average is calculated from the predicted values:

$$\widehat{y(\mathcal{S}_1, \mathcal{S}_0)}(2) = \frac{1}{N} \sum_{i \in \Omega} \widehat{m_2(\pi_i^1(2))} = 4.93$$

Finally, using the Bayesian inferential framework, the 90% posterior interval of $y\widehat{(\mathcal{S}_1, \mathcal{S}_0)}(2)$ is calculated to be the interval between 4.45 and 5.41.

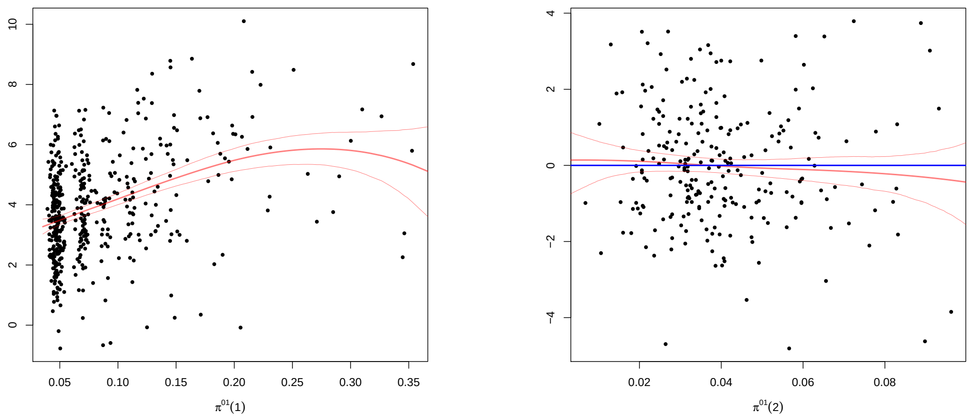
Sample Calculation for Indirect Exposure Condition

The calculation for the indirect exposure condition is a bit more tricky. First, notice that indirect exposure, by definition, cannot be observed if $c = 0$; thus, the estimator is only considered with respect to $c = 1, 2$. Second, the probability is at the level of the dyad, not the unit. However, since the data drawn from complete randomization, the spline model will use probabilities at the level of the unit i by summing all the probabilities for which i receives indirect exposure from exactly one neighbor in the network [since the dyad probabilities are nearly identical]. If the analyst wishes to consider differing probabilities at the dyad level, she is suggests to fit a two-level model with the higher level modeling probabilities at the unit level, and the lower level modeling dyadic probabilities within a neighborhood of the unit. Thus, in this discussion, the probability that unit i receives indirect exposure from exactly one neighbor under the restriction c will be denoted as $\pi_i^{01}(c)$.

Let $\mathbf{y}^{01}(c)$ denote the vector of directly treated units observed under restriction c with individual observations $y_i^{01}(c)$, and let $\widehat{m_c}$ denote the estimated mean function of $y_i^{01}(c)$ resulting from the TPRS fit with respect to assignment probability to direct treatment, $\pi_i^{01}(c)$. The estimation strategy follows the steps above beginning with $c = 1$, substituting $y^{01}(c)$ for $y^1(c)$ and $\pi_i^{01}(c)$ for $\pi_i^1(c)$. In this simulation, 20% of a unit's treatment effect spills over to neighbors. The true value of the average indirect exposure condition is 3.59.

Figure 3.7(a) demonstrates the fit of the TPRS to the indirect exposure outcome, $\widehat{m_1}$, with 90% posterior intervals for the restriction $c = 1$. The residual $\varepsilon_i^{01}(2) = y_i^{01}(2) - m_1\widehat{(\pi_i^{01}(1))}$ is calculated, and the TPRS is fit to the residual to obtain the function $\widehat{\eta_2}$, which is depicted in figure 3.7(b) with 90% posterior intervals. As before, the mean function is calculated as $\widehat{m_2} = \widehat{m_1} + \widehat{\eta_2}$. Once again, figure 3.7(b) suggests convergence to the zero function. The point estimate at $c = 2$ is 3.74 with the 90% interval 3.54 to 3.94.

Figure 3.7: Fitted Splines for Indirect Exposure Condition



(a) Fit to Outcome under $c = 1$

(b) Fit to Residual under $c = 2$

A Note about Convergence

While likelihood ratio tests have been developed for this penalized spline setting (Crainiceanu et al., 2005a), these tests may be of low power or more difficult to implement. It is imperative to develop high powered tests in this setting in order to make proper inferences about convergence. In particular, weak tests will create too lax a standard for convergence. The logic used here to determine convergence is straightforward. Once a model has been fit to the residual in the manner described in this section, one wants to test whether the predicted function has converged pointwise to the zero function.

In order to this, it is assumed that $\varepsilon_i \sim N(\widehat{\eta}_c, \sigma^2_{\varepsilon i})$ are drawn independently from the posterior, and 90% posterior intervals are analyzed pointwise to see if they bracket 0. If 90% of these posterior intervals bracket 0, then the estimator is said to converge. Of course, other standards may be used in this process. More testing is required to understand the properties of this form of testing.

3.5.5 Simulation

This subsection assesses the proposed TPRS estimator against more common approaches to conditioning on the probability of assignment, namely the inverse-probably weighted (IPW) estimators. Two other estimators are compared here, the Horvitz-Thompson (Horvitz and Thompson, 1952) and Hájek (Hájek, 1964) estimators. As is well-known, while these IPW estimators display desirable properties in terms of bias and consistency, they often perform poorly in terms of efficiency (Basu, 1971). In this spillover setting, the shrinking sample size due to distance restrictions makes efficiency all the more important.

For a given treatment assignment vector, $\mathbf{d}^*$, and indicator function I, the Horvitz-Thompson estimator, conditional on the distance restriction c, for the expected value of a treatment condition is given by:

$$y^{\widehat{HT}(\mathcal{S},\mathcal{S}'})_w(c) = \frac{\sum_{i\in\Omega}\sum_{\mathbf{d}\in\mathbf{S}_i}\frac{I(\tilde{Q}_c^i(\mathbf{d}))}{\pi_{y_i(\mathbf{d})}(c)}|\mathbf{S}_i'|w_i y_i(\mathbf{d}^*)}{\sum_{i\in\Omega}|\mathbf{S}_i||\mathbf{S}_i'|w_i} \tag{3.5.11}$$

For a given treatment assignment vector, $\mathbf{d}^*$, the Hájek estimator, conditional on the distance restriction c, for the expected value of a treatment condition is given by:

$$y^{\widehat{Haj}(\mathcal{S},\mathcal{S}'})_w(c) = \frac{\sum_{i\in\Omega}\sum_{\mathbf{d}\in\mathbf{S}_i}\frac{I(\tilde{Q}_c^i(\mathbf{d}))}{\pi_{y_i(\mathbf{d})}(c)}|\mathbf{S}_i'|w_i y_i(\mathbf{d}^*)}{\sum_{i\in\Omega}\sum_{\mathbf{d}\in\mathbf{S}_i}\frac{I(\tilde{Q}_c^i(\mathbf{d}))}{\pi_{y_i(\mathbf{d})}(c)}|\mathbf{S}_i'|w_i} \tag{3.5.12}$$

The Hájek estimator is known to be much more efficient than the Horvitz-Thompson, but if probability of assignment is accurate, the Horvitz-Thompson estimator is guaranteed to be unbiased and the Hájek estimator is guaranteed to be consistent.

To assess the relative performance of Horvitz-Thompson, Hájek, and TPRS estimators in this setting, 1000 simulations from the data-generating process in section 5.2 with $\phi_j = 0.6$ for the direct treatment condition, and 250 simulations with $\phi_j = 0.2$ for the indirect exposure condition. The results are shown in table **??**

In both cases, once $c = 2$ is reached, the TPRS estimator shows a 30% improvement in root-mean square (RMSE), which is quite significant. Thus, at higher distance restrictions, precisely where power becomes more of the a problem, the TPRS significantly outperforms the IPW estimators. At $c = 1$ the TPRS and Hájek estimator display similar performance, and the Horvitz-Thompson estimator is always significantly more inefficient and largely im-

		$c=1$	$c=2$
	truth	3.59	3.59
HT	mean	3.70	3.77
	RMSE	(0.49)	(1.30)
Hájek	mean	3.64	3.60
	RMSE	(0.12)	(0.18)
TPRS	mean	3.64	3.59
	RMSE	(0.12)	(0.13)

Table 3.2: Indirect - $y(\widehat{\mathcal{S}_{01}, \mathcal{S}_0})$

		$c=0$	$c=1$	$c=2$
	truth	4.95	4.95	4.95
HT	mean	5.31	4.95	4.96
	RMSE	(0.47)	(0.37)	(1.12)
Hájek	mean	5.31	4.96	4.95
	RMSE	(0.47)	(0.21)	(0.43)
TPRS	mean	5.31	4.93	4.92
	RMSE	(0.47)	(0.20)	(0.30)

Table 3.3: Direct - $y(\widehat{\mathcal{S}_1, \mathcal{S}_0})$

practicable in this setting. The provides strong evidence for use of the TPRS estimator. It should also be noted that the data were not drawn from a "smooth" distribution since they were drawn with respect to the degree of nodes. This suggests that even in settings where splines are not supposed to perform as well, they may still perform significantly better than IPW estimators.

Assessing Convergence

The procedure described to deduce convergence was implemented in each simulation. For the directly treated outcome, the probability of reaching convergence according to the 90% rule at $c = 1$ was 49% and at $c = 2$ was 94%, and for the indirect exposure outcome, the probability of reaching convergence at $c = 2$ was 79%. This suggests that even under relatively strong spillover biases (60% of the treatment effect in the direct case, and 20% in the indirect case), the method converges the vast majority time relatively quickly (within 2 distance restrictions).

In any sort of power analysis, the probability of deducing convergence should also be calculated. In short, in the design phase, the probability of deducing convergence, the likelihood of detecting an effect, and the likely bias should all be calculated.

3.6 Conclusion

This paper has a provided a self-contained statistical inferential approach to the estimation of causal quantities of interest in randomized experiments under spillovers and the evaluation of estimators for this purpose.

This paper provides a technique to recover causal quantities of interest without specifying an underlying stochastic process for spillovers or a priori knowledge of precisely which units share spillovers over the network. Rather, estimation is conducted with respect to a social distance measure, which is often a fairly intuitive measure (e.g., the network distance). This technique is particularly useful for detection of biases under spillovers (which is not possible in the a priori knowledge approach) and provides a more objective approach to the estimation of quantities of interest under spillovers.

One of the biggest complicating factors for causal analysis under spillovers is that the observed outcome for any unit is a function of the entire vector of treatment assignments. This paper simplifies the problem by constructing potential outcomes, treatment conditions, and quantities of interest from reference assignments. Furthermore, it is shown that experiments are well-suited to recovering "average exposure effects," which isolate the average effects of a specified type of treatment exposure over the population; these quantities are particularly of theoretical significance in analyses where a portion of the entire population is treated, perhaps due to a fixed budget. This construction allows for a full inferential framework which assesses the bias and efficiency of various estimators.

Finally, this paper proposes a Bayesian TPRS estimator and posterior inference to analyze spillovers. It is shown that TPRS provides significant gains in efficiency vis-a-vis classic IPW estimators. Inference from the posteriors is straightforward in this setting and avoids the difficulties associated with inconsistency in estimation of variance parameters in a frequentist setting. Furthermore, a simple comparison of posteriors from successive distance restrictions yields a natural test of convergence for the proposed estimator.

There are several directions for future work on this topic; two are highlighted here. As was discussed earlier, experimental design in this setting can become quite complicated. Unlike classical settings, where power is simply a function of the number treated and untreated units, the calculations are far more complicated under spillovers. The extent of complex spillovers increase as the number of units treated increases, and if too many units are treated for a given number of nodes, it may be difficult to isolate treatment conditions of interest. Furthermore, the researcher will typically have to move beyond complete randomization in order to maximize the number of units over which inferences may be made. These issues are discussed in more detail in Coppock and Sircar (2014).

At a theoretical level, the estimator will converge to true value irrespective of the social distance measure used. In practice, however, different measures will yield different estimates. In fact, it may be a good idea to estimate the model with a few natural distance measures to demonstrate the robustness of the results. But, this also yields concerns. Which estimates should the analyst believe? Intuitively, it makes sense to select the measure that yields convergence with the greatest amount of power and one may look at the sample size in each treatment condition induced by particular distance restrictions. In the case of the direct treatment effect, one may check which estimator causes the most movement from the naïve estimate.

The methods and approach discussed in this paper constitute a new, powerful way to view spillovers between units in a randomized experimental context. This framework can shed light on precisely what is being estimated when spillovers occur and the relative quality of inference of various estimators. Most importantly, however, this paper shows that estimation of causal quantities of interest is feasible with reasonable, intuitive assumptions.

Appendix A: Existing Frameworks as Special Cases

Known Structure of Spillovers

If SUTVA holds, then IPW estimator provides a consistent/unbiased estimate of the quantity of interest (direct treatment effect) under each distance restriction. To see this, notice that non-interference assumption induced by each value of c is necessarily a non-interference partition. The result follows from section 3.4. Thus, the naive estimate (with $c = 0$) yields an unbiased/consistent estimate of the quantity of interest.

Double Randomization

Now consider a blocked design with double randomization, where spillover occur within blocks but not across them. This exposition follows that of Tchetgen Tchetgen and VanderWeele (2010) closely.

Let the sample of experimental units be denoted by Ω, nested in M blocks, with a total sample size of N. The sample size of block $b_k \in \{b_1, \dots, b_M\}$ is denoted as N_k. For unit $i \in \Omega$, the block containing i will be denoted by $b(i)$.

In double randomization, first a subset of M_1 blocks are randomly chosen to receive the condition α_1, with the remaining blocks receiving α_0. Typically, α_0 and α_1, $0 < \alpha_0 < \alpha_1 < 1$, correspond to the fraction of units receiving treatment in the corresponding block with the condition assigned to block b_k denoted as $\alpha(b_k)$, so the condition received by the block containing unit i will be denoted as $\alpha(b(i))$. Treatment to each unit will be binary $\mathbf{d}_i \in \{0, 1\}$, i.e. untreated or treated.

There standard quantities of interest, the total effect and the indirect effect, will be discussed. The *total effect* is the effect of changing the treatment condition from being in a block with α_0 while untreated to a block with α_1 while treated. The *indirect effect* is the effect of changing the treatment condition from being in a block with α_0 while untreated to a block with α_1 while untreated.

To keep the calculations clean, it will be assumed that $\alpha_0 N_k$ and $\alpha_1 N_k$ are integer-valued for all blocks k. There are $\binom{M}{M_1}$ ways of selecting the conditions for blocks, and within each block k there are $\binom{N_k}{\alpha(b_k)N_k}$ ways of implementing the condition $\alpha(b_k)$.

At the level of the experimental unit, there are 3 treatment conditions of interest: 1) $\mathcal{S}_{\alpha_1 1}$ – direct treatment in a block receiving α_1; 2) $\mathcal{S}_{\alpha_1 0}$ – no treatment received in a block receiving α_1; and 3) $\mathcal{S}_{\alpha_0 0}$ – no treatment received in a block receiving α_0. Consider a binary treatment assignment vector ($\mathbf{d}_i \in \{0,1\}$ for all $i \in \Omega$). The corresponding treatment conditions for an experimental unit i may be written as a set of assignment vectors, $\mathbf{d}$:

- *Treated in a block with* α_1 –

$$\mathbf{S}_{\alpha_1 1i} = \left\{ \mathbf{d} \;\middle|\; \mathbf{d}_i = 1, \sum_{\{j|b(i)=b(j)\}} \mathbf{d}_j = \alpha_1 N_{b(i)}, \sum_{\{j|b(i)\neq b(j)\}} \mathbf{d}_j = 0 \right\}$$

- *Untreated in a block with* α_1 –

$$\mathbf{S}_{\alpha_1 0i} = \left\{ \mathbf{d} \;\middle|\; \mathbf{d}_i = 0, \sum_{\{j|b(i)=b(j)\}} \mathbf{d}_j = \alpha_1 N_{b(i)}, \sum_{\{j|b(i)\neq b(j)\}} \mathbf{d}_j = 0 \right\}$$

- *Untreated in a block with* α_0 –

$$\mathbf{S}_{\alpha_0 0i} = \left\{ \mathbf{d} \;\middle|\; \mathbf{d}_i = 0, \sum_{\{j|b(i)=b(j)\}} \mathbf{d}_j = \alpha_0 N_{b(i)}, \sum_{\{j|b(i)\neq b(j)\}} \mathbf{d}_j = 0 \right\}$$

Since the quantity of interest makes inferences about the effect of treated units in the block, the reference assignments for unit i correspond to assignment vectors where only units in block containing i are treated. In the analysis of such experiments, typically the cluster averages for the quantities of interest are calculated. The quantities of interest, total effect (TE) and indirect effect (IE), may now be written as:

$$TE = \frac{1}{M}\sum_{k=1}^{M}\frac{1}{N_k}\sum_{i\in b_k}\frac{1}{|\mathbf{S}_{\alpha_1 1i}||\mathbf{S}_{\alpha_0 0i}|}\sum_{\mathbf{d}\in\mathbf{S}_{\alpha_1 1i}}\sum_{\mathbf{d}'\in\mathbf{S}_{\alpha_0 0i}} y_i(\mathbf{d}) - y_i(\mathbf{d}') \tag{3.6.1}$$

$$= \frac{1}{M}\sum_{k=1}^{M}\frac{1}{N_k}\sum_{i\in b_k}\frac{1}{\binom{N_k-1}{\alpha_1 N_k-1}}\sum_{\mathbf{d}\in\mathbf{S}_{\alpha_1 1i}} y_i(\mathbf{d}) - \frac{1}{M}\sum_{k=1}^{M}\frac{1}{N_k}\sum_{i\in b_k}\frac{1}{\binom{N_k-1}{\alpha_0 N_k}}\sum_{\mathbf{d}'\in\mathbf{S}_{\alpha_0 0i}} y_i(\mathbf{d}')$$

$$IE = \frac{1}{M}\sum_{k=1}^{M}\frac{1}{N_k}\sum_{i\in b_k}\frac{1}{|\mathbf{S}_{\alpha_1 0i}||\mathbf{S}_{\alpha_0 0i}|}\sum_{\mathbf{d}\in\mathbf{S}_{\alpha_1 0i}}\sum_{\mathbf{d}'\in\mathbf{S}_{\alpha_0 0i}} y_i(\mathbf{d}) - y_i(\mathbf{d}') \tag{3.6.2}$$

$$= \frac{1}{M}\sum_{k=1}^{M}\frac{1}{N_k}\sum_{i\in b_k}\frac{1}{\binom{N_k-1}{\alpha_1 N_k}}\sum_{\mathbf{d}\in\mathbf{S}_{\alpha_1 0i}} y_i(\mathbf{d}) - \frac{1}{M}\sum_{k=1}^{M}\frac{1}{N_k}\sum_{i\in b_k}\frac{1}{\binom{N_k-1}{\alpha_0 N_k}}\sum_{\mathbf{d}'\in\mathbf{S}_{\alpha_0 0i}} y_i(\mathbf{d}')$$

The quantities of interest follow from the fact that $|\mathbf{S}_{\alpha_1 1i}| = \binom{N_{b(i)}-1}{\alpha_1 N_{b(i)}-1}$, $|\mathbf{S}_{\alpha_1 0i}| = \binom{N_{b(i)}-1}{\alpha_1 N_{b(i)}}$, and $|\mathbf{S}_{\alpha_0 0i}| = \binom{N_{b(i)}-1}{\alpha_0 N_{b(i)}}$. The weights for the quantity TE are $w_i = \frac{1}{N_{b(i)}|\mathbf{S}_{\alpha_1 1i}||\mathbf{S}_{\alpha_0 0i}|}$, and the weights for the quantity IE are $w_i = \frac{1}{N_{b(i)}|\mathbf{S}_{\alpha_1 0i}||\mathbf{S}_{\alpha_0 0i}|}$.

Notice, once again, that these quantities of interest are formed without making any assumptions about the structure of spillovers between the units. Now, consider a class of distance measures that fit the following criteria:

$$\rho(i,j)\ s.t.\ \begin{cases} \rho(i,j) = 0 & \text{if and only if } i = j \\ \rho(i,j) \le \underline{c} & \text{if and only if } b(i) = b(j) \\ \rho(i,j) > \underline{c} & \text{if and only if } b(i) \ne b(j) \end{cases}$$

This implies that the blocks are induced by a certain class of distance measures, but the blocks must be formed before experiment since they affect the randomization procedure. Based upon the distance measure, the probability of inclusion under a distance restriction, c, can be written easily through conditional probabilities:

$$\pi_{y_i(\mathbf{d})}(c) = P(\alpha(b(i))) * P(\mathbf{S}_{\nu i}(c) \mid \alpha(b(i))) * P(\mathbf{d} \mid \mathbf{S}_{\nu i}(c), \alpha(b(i))), \quad \nu \in \{\alpha_0 0, \alpha_1 0, \alpha_1 1\} \tag{3.6.3}$$

Under the assumption of no spillovers across blocks, the probability $\pi_{y_i(\mathbf{d})}(\underline{c})$ is selected. Typically, blocks are selected equal probability to receive α_0 or α_1, and $\alpha(b_k)N_k$ units receive treatment in b_k through complete randomization with equal probabilities of inclusion.

If unit i is in block b_k receiving treatment, where block has $\alpha_1 N_k$ units treated, then the corresponding probability of inclusion is given by:

$$\pi_{y_i(\mathbf{d})}(\underline{c}) = \frac{\binom{M-1}{M_1-1}}{\binom{M}{M_1}} * \frac{\binom{N_k-1}{\alpha_1 N_k-1}}{\binom{N_k}{\alpha_1 N_k}} * \frac{1}{\binom{N_k-1}{\alpha_1 N_k-1}} = \frac{M_1\alpha_1}{M * \binom{N_k-1}{\alpha_1 N_k-1}} \tag{3.6.4}$$

If unit i is in block b_k receiving no treatment, where block has $\alpha_1 N_k$ units treated:

$$\pi_{y_i(\mathbf{d})}(\underline{c}) = \frac{\binom{M-1}{M_1-1}}{\binom{M}{M_1}} * \frac{\binom{N_k-1}{\alpha_1 N_k}}{\binom{N_k}{\alpha_1 N_k}} * \frac{1}{\binom{N_k-1}{\alpha_1 N_k}} = \frac{M_1(1-\alpha_1)}{M * \binom{N_k-1}{\alpha_1 N_k}} \tag{3.6.5}$$

If unit i is in block b_k receiving no treatment, where block has $\alpha_0 N_k$ units treated:

$$\pi_{y_i(\mathbf{d})}(\underline{c}) = \frac{\binom{M-1}{M-M_1-1}}{\binom{M}{M-M_1}} * \frac{\binom{N_k-1}{\alpha_0 N_k}}{\binom{N_k}{\alpha_0 N_k}} * \frac{1}{\binom{N_k-1}{\alpha_0 N_k}} = \frac{(M-M_1)(1-\alpha_0)}{M * \binom{N_k-1}{\alpha_0 N_k}} \tag{3.6.6}$$

Consider an observed assignment vector of $\mathbf{d}^*$. Plugging the probabilities induced by the restriction $\underline{c}$ into the IPW estimators yields the usual estimators of differences in the average block means for the respective conditions:

$$\widehat{TE}(\underline{c}) = \frac{1}{M_1}\sum_{k=1}^{M}\frac{1}{\alpha_1 N_k}\sum_{i\in b_k} y_i(\mathbf{d}^*)I(\mathbf{d}_i^* = 1, \alpha(b_k) = \alpha_1) \tag{3.6.7}$$

$$-\frac{1}{M-M_1}\sum_{k=1}^{M}\frac{1}{(1-\alpha_0)N_k}\sum_{i\in b_k} y_i(\mathbf{d}^*)I(\mathbf{d}_i^* = 0, \alpha(b_k) = \alpha_0)$$

$$\widehat{IE}(\underline{c}) = \frac{1}{M_1}\sum_{k=1}^{M}\frac{1}{\alpha_1 N_k}\sum_{i\in b_k} y_i(\mathbf{d}^*)I(\mathbf{d}_i^* = 0, \alpha(b_k) = \alpha_1) \tag{3.6.8}$$

$$-\frac{1}{M-M_1}\sum_{k=1}^{M}\frac{1}{(1-\alpha_0)N_k}\sum_{i\in b_k} y_i(\mathbf{d}^*)I(\mathbf{d}_i^* = 0, \alpha(b_k) = \alpha_0)$$

As before the function I denotes an indicator function that takes the value 1 if the condition is met and 0 otherwise.

The double randomization design is dependent upon a known structure of blocks a priori, where spillovers occur within blocks and not across them. This suggests that all plausible distance measures must yield the exact same block structure.

Once again this framework defines the quantity of interest independently of any interference assumptions. In doing so, it is possible to provide some guidance on extending the double randomization framework. Suppose the researcher is concerned there may be spillovers between two blocks, b_j and b_k, for which the same condition the same condition is generated, $\alpha(b_j) = \alpha(b_k) = \alpha_1$. In the final analysis, the researcher decides to run the analysis in two ways, under the standard double randomization estimator and by combining blocks j and k (properly accounting for the probability that $\alpha(j) = \alpha(k) = \alpha_1$).

Even if the researcher gets different results, she will be unable to determine whether there actually were or were not spillovers between block b_j and b_k, and thus unable to select to a valid result. To see why, notice that the researcher has unwittingly changed the underlying quantity of interest. When the researcher combined blocks, she failed to notice that the estimator for combined blocks assumes that there are $\binom{N_j+N_k}{\alpha_1(N_j+N_k)}$ ways of selecting the condition α_1 at the combined block of b_j and b_k. But this is not true if the blocks have been combined; there are only $\binom{N_j}{\alpha_1 N_j} + \binom{N_k}{\alpha_1 N_k}$ ways of selecting α_1 jointly and thus certain reference assignments have been omitted in this quantity.

This is clearly a very important problem in double randomization; one cannot simply assume that there are no spillovers across blocks (when there may in fact be spillovers). Remember, however, that the blocks have been induced by a distance measure, ρ. Using the usual procedure of considering more stringent restrictions under ρ, $c > \underline{c}$, one can arrive at a defensible estimate even if there is a fear that there may be spillovers across blocks. Notice that the procedure has nothing to do with combining or dropping blocks, but rather taking subsets of blocks for which the outcomes of interest are cleanly observed.

Chapter 4

A Tale of Two Villages: Kinship Networks and Preference Formation in Rural India

Abstract

This study investigates the effect of kinship networks on vote choice and issue preferences over an electoral campaign in rural India. The study analyzes data collected on political preferences and kinship networks in two villages just before and after the campaign period during the 2011 Assembly election in West Bengal. The paper finds very strong kinship network effects on changes in political opinions and vote choice over a campaign. It is argued that this is due to information pooling, political discussion and explicit coordination of political behavior within the family, which results from the codependence between members of a family. Based on eight months of direct observation around the election, this paper provides strong qualitative evidence for the proposed mechanisms. Furthermore, using a network autoregressive lag model, data on vote choice and ideal point estimation, the paper provides fine grained quantitative information on the role of kinship networks in changing vote choice and issue preferences.

4.1 Introduction and Motivation

The study of voting behavior in a democratic developing country context is dominated by accounts of clientelism and patronage (Chandra, 2004; Stokes, 2005; Posner, 2005; Kitschelt and Wilkinson, 2007; Lust, 2009). Clientelism and patronage are, however, equally present in non-democratic contexts (Bates, 1981; Anderson, 1987), so the focus of this literature has been to describe how clientelism and patronage can persist in democratic settings. While clientelism and patronage are undoubtedly present, less attention has been given to features of democratic practice unique to developing world democracies, even as the number of stable democratic developing countries is growing. Indeed, there is little literature in democratic developing countries on one of the most fundamental questions in democratic practice: how do voters form preferences?[1]

This paper endeavors to provide an account of how voters in a democratic developing country context form their political preferences. This study focuses on two villages just before and after the electoral campaign period during the 2011 Assembly election in the Indian state of West Bengal. Democratic developing country contexts are often characterized by weak state capacity and low information about politically salient matters. The campaign period around an election constitutes a democratic moment where voters are required to piece together information from disparate sources on the ability of a candidate to govern.

Kinship networks aid in this task by providing information pooling, political discussion and explicit coordination of political behavior in order to develop and update political preferences. This political coordination across a kinship network is nested within the larger role of the kinship network in fostering cooperation to mitigate physical and economic risks in rural contexts. A voter by herself, with little information about politically salient issues, is vulnerable to manipulation from political actors. However, by hooking into her kinship network, and pooling and reasoning over salient information, the voter is able to make an

[1] I thank S. Chandrasekhar, Devesh Kapur, Macartan Humphreys, Sripad Motiram, Armando Razo and Steven Wilkinson for helpful comments, as well as participants at the SSDS seminar at Columbia University, Indira Gandhi Institute for Developmen Research, MPSA 2013 and PolNet 2013. Research was funded by grants from the Applied Statistics Center and Center for the Study of Development Strategies, both housed at Columbia University.

informed decision and magnify the impact of her decision through coordination over the kinship network. It is, therefore, the kinship network that engenders the independence of the voter.

In order to demonstrate these claims, the paper discusses fine-grained qualitative and quantitative evidence on the role of kinship networks in changing vote choice and issue preferences over the electoral campaign. Using eight months of qualitative field research around a single election in two villages, this paper provides detailed information on the characteristics of kinship networks, their structural connection to existing political preferences, and their role in changing political preferences. The data, using information on vote choice and issues preferences combined with ideal point estimation and network autoregressive models, provide quantitative justification for the claims.

4.1.1 Democracy in the Developing World

Today, democracies comprise a significant proportion of developing world countries. Beginning with democratic transitions in Southern Europe, and Latin America in the 1970s and 1980s, as well as transitions in Africa, Southeast/East Asia and Eastern Europe in the 1990s and the 2000s, the expansion of democracy is the fruit of what has collectively been referred to as the "third wave of democracy" (Huntington, 1991). This expansion spawned a literature on democratic transitions and democratic consolidation in the developing world (Stepan and Linz, 1996; Przeworski et al., 2000). While this is undoubtedly an important literature, this paper asks a different question. After democratic consolidation has taken place in a developing world country, what does democracy look like?

This paper focuses on a rural Indian setting. On the whole, India is more than two-thirds rural (Census of India, 2011), making it one of the most agrarian-based democracies in the world. Many developing world democracies, mostly recently consolidated, display very large rural populations, such as countries in Africa, Central America, as well as South and Southeast Asia. Like India, many of these countries also exhibit weaker state capacity (Migdal, 1988), which hampers their abilities to properly enact policies without bureaucratic or political manipulation. At the same time, Indian democracy is a strongly consolidated democracy; it is just one of 33 countries (and by far the poorest and least literate of such countries) that has been continuously democratic since 1977 (Lijphart, 2006). This makes India a particularly good place to understand the longer run aspects of voter behavior in burgeoning developing world democracies.

4.1.2 Studying the Indian Voter

Much of the political science literature dealing with political behavior or voter behavior in a democratic developing country context focuses on patronage, clientelism or vote-buying.[2] While this is undeniably a large part of politician-voter interaction in the developing world, it mutes the democratic deepening that is taking place in many of these countries over time.

In the context of India, many of these claims are an outgrowth of the literature in the 1960s on the so-called "Congress system,"[3] which focused attention on how the Congress Party, the party that controlled national government following Indian independence, co-opted elites and manufactured a strong patronage system. These elite-centric and party-centric arguments diminished the importance of the Indian voter and little effort was put into understanding how the average Indian voter forms political preferences. However, more recent literature has focused on the democratic deepening of India. Stepan et al. (2011) report robust support for democracy and democratic principles in India.[4] Many studies have shown an increase in formal political actors from lower classes and castes, signifying a breakdown of elite domination.[5] In fact, while essentially characterized by single party rule from 1947-

[2]Patronage or clientelism fundamentally involves "imbalance in exchange...which expresses and reflects the disparity in...relative wealth, power, and status." (Scott, 1972, p. 93) Kitschelt and Wilkinson (2007) edited an entire volume on the topic, with examples from Latin America, Africa, and South Asia. Stokes (2005) discuss how political parties monitor voters in Argentina to enact a system of clientelism. Chandra (2004) refers to India as a "patronage democracy" and derives a model by which caste groups vote for co-ethnic politicians in exchange for patronage.

[3]The term "Congress system" was first coined in Kothari (1964) and was further developed in a comprehensive study by Weiner (1967).

[4]Furthermore, Banerjee (2011), also based upon anthropological work in West Bengal, shows that elections have taken on increased cultural significance, even displaying sacred and ritualistic elements.

[5]Krishna (2002) and Manor (2000) have demonstrated the rise of a new class of brokers, through whom villagers can access public goods and services, whose viability relies on the ability to deliver goods and not social status. Jaffrelot and Kumar (2009) and Michelutti (2009) have chronicled the "subalternization" of Indian politics, whereby lower castes are entering the formal political arena in greater numbers.

1977, India has, more recently, tended to be characterized by party/candidate alternation and anti-incumbency (Linden, 2004). This alternation and belief in democracy would seem to militate against the patronage/clientelism view. As India's political culture has become more democratic, and traditional structures of co-optation have deteriorated, it has become important to understand how the average Indian voter forms political preferences.

4.1.3 Argument

Using data from two villages in the Indian state of West Bengal, this paper demonstrates that kinship networks are used to pool political salient information and to explicitly discuss and coordinate political behavior to generate political preferences of voters. In particular, in order to make an informed choice under an environment with weaker state capacity, the voter requires information about various candidate characteristics due to the party's or candidate's capacity to manipulate state resources and powers. This information can affect the voter in two ways. First, she may attempt vote for the winning candidate in order to opt in to a clientelistic scheme. Alternatively, the voter may seek detailed information on a candidate's qualifications since weak state capacity increases the salience of a politician's personal skill and capacity in performing his duties.

The electoral campaign provides information on many of these factors and the kinship network acts to pool information gleaned by various family members. Kinship networks are also used to explicitly coordinate political and voting behavior. This coordination may be due to an explicit desire for the family to demonstrate its support for the winning party, or it may simply be due to the larger structure of codependence over a kinship network. The paper finds that in addition to changing vote choice, kinship networks affect ideological preferences, suggesting an information role for kinship networks that move beyond patronage. Importantly, these kinship structures act to mitigate the costs associated with updating personal political preferences, costs which are particularly acute in a developing democratic context.

This paper extends classic theories about social influence, e.g., the Columbia School, to an expanded role for families in information pooling and political coordination. This is important because it shows that the underlying social structure in a developing country helps mitigate difficulties caused by lower levels of economic and institutional development in a democratic context. The information pooling function of kinship networks helps to piece together politically salient information about candidates, parties and policies, which would otherwise be difficult for voters to access. The discussion and coordination function of kinship networks allows voters to reason over complicated and disparate pieces of information to make informed choices about the election in an environment with low levels of literacy and where political leaders can manipulate state resources and powers. Fundamentally, this ability to coordinate behavior, both vote choice and ideological preferences, over kinship networks demonstrates that, despite coercive pressures from political actors above, voters can act independently due to a family's capacity to provide a forum insulated from other political forces. Paradoxically, the independence of the voter results from an individual giving away her personal agency to the kinship network at large.

4.1.4 Layout

Using extensive fieldwork and data on kinship networks, this paper provides fine-grained qualitative and quantitative information on the role of kinship networks in changing vote choice and issue preferences. The empirical results rest on three claims which are shown in detail:

1. The campaign period has an effect on vote choice and issue preferences
2. There are strong kinship network effects on vote and opinion change over the campaign
3. Kinship network effects can be largely attributed to political discussion and coordination, even when controlling for other prominent explanations

In short, this paper juxtaposes close qualitative observation in two villages and data collection in the same villages to demonstrate strong evidence for its propositions. This affords the opportunity to collect detailed village-level data, as well as providing meaningful explanation of data trends.

Section 2 lays out the theory of how family coordination and discussion affect political behavior. Section 3 discusses the qualitative evidence and study design. Section 4 demonstrates that vote choice and issues preferences change over the campaign. Section 5 demonstrates strong kinship network effects in changes in vote choice and political opinions. Section 6 demonstrates that kinship network effects are largely due to political discussion and coordination of political behavior, even when controlling for other prominent outcomes. Section 7 concludes the paper, discussing larger implications for developing world democracies as well as hypotheses to be tested in the future.

4.2 Theory

The modern social network approach to political behavior has its origins in the so-called "Columbia School" of sociologists, who studied American voting behavior in the middle of the 20^{th} century. The corresponding studies argued that vote choice and political opinions were largely a function of one's own personal network. Much like the present study, these claims were substantiated by survey research at the community level in Erie County, Pennsylvania (Lazarsfeld et al., 1944) and Elmira, New York (Berelson et al., 1954). The Columbia School also noticed the prominent role occupied by kinship networks, viewing them as the most important drivers of political identities. At the same time, they argued that individuals generally choose to seek out information that reinforces their views; as such, the Columbia School viewed campaigns and media as having little effect on actual political opinions. They also held a dim view of overly individualistic theories where individuals independently made strategic, rational voting decisions given their states of knowledge.

While acknowledging the role of kinship, Campbell et al. (1960) criticized the Columbia School as too focused on social influences on political opinions. They argued that individuals are "socialized" into a particular partisan identity early in life, usually through parents, and these early partisan identities had long-lasting impact on subsequent political beliefs. Once again, these theories found little room for the impact of campaigns on political opinions. The rationalist school of thought more generally criticized the Columbia School for social determinism but also found a role for campaigns and media. They argued that one's friends and family may act as "information shortcuts" to process the political information generated in a campaign, after which voters make fully rational decisions (Downs, 1957; Popkin, 1994; Lupia and McCubbins, 1998). This constellation of theories omits one important scenario, that campaigns may have an effect of political opinions more generally but decisions are not made by independent individuals. This lacuna is not surprising given that these are not intended to explain political behavior in a developing world democratic context like India.

This paper argues that voters discuss politics and coordinate vote choice through kinship networks. Kinship networks affect preferences by acting as a vessel to pool and discuss relevant political information and as an implement to coordinate voting behavior. Much like the original Columbia School, it is argued overly individualistic theories do not appropriately capture a voter's decision to vote or formation political opinions. However, this argument expands role of kinship beyond what was initially envisioned by the Columbia School. Apart from social influence on opinions, this paper views kinship as instrumental in strategic coordination of political behavior.

This study finds that campaigns have an impact on both vote choice and political opinions, with such effects flowing through kinship networks. This provides a point of departure from the Columbia School. The campaign provides salient political information about which party is likely to win the election and policy positions, as well as information about a candidate's ability to deliver, protect, and govern. The campaign period is a natural time for voters to update their political preferences. Since updating preferences requires political education, it is a costly endeavor. The campaign period is a natural time to update preferences due to an increased flow of political information and the incentives to update preferences before the upcoming election.

These impacts are magnified in a weak state environment because elected politicians are more able to condition benefits and protection on partisan support without the encumbrances of the formal state, and, even when not engaging in clientelism, rely on personal skill and capacity in delivering benefits. Furthermore, certain candidate characteristics, such as ethnicity and criminality, may actually provide a credible signal of a candidate's ability to deliver benefits or protect the population. Typically, candidates are announced at the start of the campaign period and often little is known about them. The value of discovering the personal characteristics of candidates, even apart from the increased expected benefit of supporting the winner, provides powerful incentives for families to pool information to update political opinions and strategize over vote choice.

4.2.1 Kinship

Defining kinship can be a difficult task. This paper puts forth a network conception of kinship, as opposed to a group-based conception of family. As Inden and Nicholas (1977) have shown, consanguinity, a standard criterion for kinship in Western societies, does not fully characterize the South Asian family. For instance, a woman who marries into a family becomes a part of that family. This cultural understanding of what constitutes a family is crucial to any analysis of kinship. At the same time, it is important distinguish between the relative distance in relationship between family members. Two women who have married into the same family are likely to be more distant than those who have spent a significant portion of their lives together, like siblings or parent/child. In the quantitative portion of this study, two individuals are linked in the kinship network if they satisfy a "nuclear relation": sibling, spouse, parent, or child. This effectively characterizes the South Asian notion of a family (since two women who have married into the same family will still be connected but more distant than two brothers), while accounting for the relative closeness of family members.

Little work exists on the importance of families for political decision-making in India or the rest of the developing world. This is all the more surprising given the importance of families in Indian society. The last National Election Survey in India (Lokniti, 2009) found that for 24.5% of respondents the views of a spouse or other family member mattered the most, even more than one's own opinion, in voting decisions.

Kinship represents the most prominent and influential social and personal network in a villager's life. Due to the traditional nature and spatial arrangement of villages, a villager typically interacts regularly with her extended family. Intra-household coordination is natural in a poorer rural context, as it is often used in employment and marriage decisions to mitigate risks from consumption shocks (Rosenzweig, 1988; Rosenzweig and Stark, 1989). The implicit assumption in this literature is that families are able to devise methods to maintain cooperative behavior among their members (Lucas and Stark, 1985). This extraordinary ability of kinship ties to maintain cooperation makes it a natural place to observe coordinated political behavior.

There is a large literature on the association between social/ethnic identity of voters and partisan preferences in developing societies.[6] In India, as in many other developing contexts, politics is coordinated at the village level through village-level political leaders and workers (Kruks-Wisner, 2011; Bussell, 2014) and identity is often too blunt an object to understand political differences and changes at the village level, which typically have a few castes and religions within them.[7] For example, while Muslim voters may, in the aggregate, lean towards a specific party in the polity, this does not imply that entire population of a fully Muslim village will vote for that party. Generally, a fully Muslim village, like any other village, will have factions supporting multiple parties. The relationship between these factions and families has been known for some time; the seminal work on factionalism in Indian villages, Lewis (1954), found that villagers "tend to equate their faction with their kinship group." At the same time, the critical role of the family in political decision-making remains understudied.

[6]The existing literature posits instrumental calculations between co-ethnic voters over patronage (Chandra, 2004), psychic rewards for voting for co-ethnics (Chandra, 2009), and elite manipulation to construct disparate ethnic "minimum winning" coalitions (Posner, 2005) as potential mechanisms to explain this association. While these may be useful mechanisms to describe politics in the aggregate, it can be difficult to apply these theories at the local level.

[7]In fact, the recent National Election Survey finds that only 5% of respondents list the opinions of caste or community leaders as mattering the most for their vote choice.

4.2.2 Influence and Coordination in Opinion Change

Campaigns and Information Pooling over the Kinship Network

The data from the United States suggest minimal effects of the campaigning on political opinions. Gelman and King (1993) find that election results can be predicted within a couple of percentage points a few months before the presidential elections, echoing findings by the Columbia School. They argue that voters only piece together salient information for their vote choice in the days before the election, but this salient information is readily apparent months before an election to analysts.

The Indian case differs from this standard. A significant amount of necessary information, from a voter's perspective, is not apparent until the campaign begins. In particular, candidates are typically announced at the start of a campaign, and candidate characteristics may matter in an election. Given the relatively low intra-party democracy of most Indian parties, there is often little information about the selected candidates. A candidate's ethnic background (Chandra, 2004) or criminal background (Vaishnav, 2012a) may serve as a credible signal of a candidate's ability to deliver benefits. A second piece of crucial information during the campaign period concerns the winnability of a candidate and a party. The major media houses will typically provide a pre-election projections for an election at this time. In a system that often displays a significant amount of volatility in vote shares, this can provide new, concrete information. Patnam (2013) demonstrates that unexpected information about the winnability of a party (through exit polls) may cause as much as a twenty point increase in the probability of voting for that party. Finally, reasoning through disparate pieces of politically salient information requires a lot of effort, perhaps much more than the American context, and many voters only undertake this effort and update their preferences during an electoral campaign. Thus, changes in beliefs may be less due to information and more due to the process of updating beliefs.

Families often play an informative role in politics, with individuals sharing information and educating each other about various issues. The literature of social influence generally takes two forms, political socialization and political discussion. The idea of political socialization, especially in regards to the political influence of a parent towards a child, has been studied extensively in American politics. It is generally argued that parents play a crucial role in inculcating particular preferences in children, the so-called direct transmission theory (Glass et al., 1986). Others have argued for more nuanced approach while allowing for the basic idea that parents inculcate political preferences (Jennings et al., 2009).

Political discussion is much more firmly associated with the Columbia School. However, looking at American elections, the Columbia School believed political discussion simply reinforced existing political positions since individuals would seek out personal networks with agreeable political positions. As such, the Columbia School did not believe campaigns or media had much effect on political opinions or change. The notion of political discussion within a kinship network as conceived in this paper differs in two major ways. First, as described above, campaigns do provide salient information to voters that are generally required for a reasoned discussion. Second, kinship is not selected and makes up the lion's share of a villager's personal network. Thus, while families may experience common family histories, members of the family vary widely in age, education, and other characteristics (unlike friendship networks) that imply that members of a family are often exposed to disparate sources of information. Kinship networks can then serve to pool a rich variety of salient political information and to act as a conduit for discussion.

Political Coordination over the Kinship Network

Information pooling and political discussion, however, does not necessarily imply families are coordinating political behavior. Why should members of the kinship network willingly give away personal agency in political decisions for coordination over a kinship network?

One of the major themes of the study of democratic behavior in India, and the developing world more generally, has been the use of voting to access protections or benefits from the state. Chandra (2004) argues that India is a "patronage democracy" where voters support a party so that the party may deliver benefits directly to its supporters through ethnic cues. Chhibber and Nooruddin (2007) argue that voters observe the state's ability to spend and direct funds in the future to make decisions about whom to support. What is common to these theories, and many others on voter behavior in these contexts, is that the state is not seen as objective arbiter of who will receive benefits and protections, and parties and local political actors may condition state benefits or protection upon political support.

Targeted benefits received from the state are rarely given at the individual level, e.g. jobs. As Vaishnav (2012b) has shown, India actually has the lowest per capita public sector employment of any of the G20 economies. Direct payoffs and individualized benefits are remarkably inefficient in the Indian political context. By contrast, many of the targeted benefits are the sort of goods that are likely to benefit an entire kinship network and beyond, such as roads or potable water (Bardhan et al., 2011). Because the entire family is likely to benefit from any good, there is an incentive for the family to coordinate its votes, especially since political actors are likely to condition family benefits upon the depth of support within the family (Bardhan et al., 2009).

Under a secret ballot electoral system, as in India, it is typically difficult to enforce political coordination between families and political actors; as Stokes (2005) has argued, this sort of coordinated relationship requires "monitoring" of political choices. While not formally monitoring per se, this coordination can be enforced through the density of social ties in the village. In a village, an individual is being constantly observed by many others, from what she says to those with whom she associates. In this high information environment, while vote choice cannot be observed, one's commitment to a party can reasonably be observed, from showing up to political rallies to regularly association with party members. If benefits are distributed with respect to *demonstrated* support for a party (Bardhan et al., 2009, 2011),[8] this level of information provides a credible mechanism to enforce cooperation within the household and condition benefits based upon political support. Nonetheless, the act of demonstrating support is a costly activity that may be undertaken by the kinship network.[9]

A second mechanism results from the ability of kinship networks to mitigate physical

[8]Using a survey 89 village across West Bengal, these survey find that nearly 70% of respondent make financial contributions to political campaigns, and 48% participate in party meetings.

[9]It is important to note that a wide-ranging analysis of the American case by **?** has found limited evidence that voters act out of self-interest.

and economic risks. As described above, often villagers display a significant amount of codependence in relations across the kinship network. At times the cost of opting in to the clientelistic scheme described above may be too costly for the kinship network. Even in such a scenario, detailed information is required about the personal skill and capacity of candidates to deliver benefits since not all benefits are distributed along clientelistic lines. In a context where family-level preferences are at a premium, which is to say that there exists a norm that all family members should have similar preferences, a coordinated political choice implies a larger bloc of votes to the candidate/party of choice. Even if the vote choice of family members is unknown to other villagers, an uncoordinated vote choice only serves cancel out the broader impact of the family since family members are voting for opposing candidates/parties. In this context, it is best for an entire family unit to coordinate its vote choice to the greatest extent possible to maximize its impact. Thus, a kinship network may choose to shelter itself from monitoring from political actors while still coordinating vote choice; it is this ability to shield from political actors, while creating a mechanism for informed political decisions, that generates independent voter behavior.

In order to demonstrate the role of political discussion and coordination in opinion change, this paper marshals three pieces of evidence:

1. Campaigns strongly affect political opinions and vote choice

2. This effect flows through kinship networks

3. The effect of kinship networks can be attributed to political discussion and coordination

4.3 Study Design and Qualitative Evidence

The study took place in two villages in the Indian state of West Bengal. West Bengal has its own unique political history. The Communist Party of India (Marxist) or CPM was, at the time, considered the most organized political party in India, and, as a continuously elected leftist party for 34 years, the party exercised very strong control over all state institutions and personal networks in West Bengal through which it distributed patronage (Mallick, 1993).

Furthermore, 68% of the population of West Bengal is rural (Census of India, 2011). This suggests that large political shifts in West Bengal are likely to be due to changes in support from the rural population. The political history of West Bengal provides another interesting reason to focus on rural voters. After the CPM came to power in 1977, it forged a strong rural base through land redistribution programs. In 1972, a law was enacted to restrict formal landholding to a maximum of 5-7 hectares (about 12.5-17.5 acres) per family based on size, which was poorly enforced. Using a combination of violent takeover of land (Ruud, 2003) and policies to grant titles to land, the CPM built its rural base. This effectively took land away from the traditional landowning class, or *zamindars*, and redistributed the land to the landless. Two policies were particularly notable in this task: 1) *operation barga*, which sought to register and dole out land to sharecroppers, or *bargadars*; and 2) a *patta* (land titling) program which gave land titles on vested lands which had often been extracted from zamindars (Bardhan and Mookherjee, 2003).

In addition to land reform, the CPM had developed a strong grassroots base, with connection to youth through campus politics and to people associated with various occupations through unionization. Yet, despite its massive advantage in organization and providing patronage, the CPM lost the control of the state in May 2011. While there were many underlying reasons for the collapse, the proximate cause was the government's decision to expropriate land in the villages of Singur and Nandigram, which set off a wave of protests and demonstrations against the government. In fact, the CPM and its allies only mustered only 63 out of 294 seats in the last state assembly election (the new ruling coalition of Congress and Trinamool Congress (TMC) received 227 seats). Given the CPM's level of organization and its ability to insert itself into personal networks, West Bengal provides a particularly interesting case in which to test the extent to which coordination over the kinship network contributed to this change.

4.3.1 Villages under Study: Ranjanpur and Chaandinagar

Two villages, Ranjanpur and Chaandinagar, were chosen with respect to the *diverse case design* (Seawright and Gerring, 2008). In particular, two villages were selected from the same electoral constituency but with very different underlying demographic characteristics. Holding the constituency constant across the study guarantees that any observed differences between the villages of study are not due to constituency-level differences. As discussed in detail below, Ranjanpur is a poorer, underdeveloped village with a Muslim population, whereas Chaandinagar is wealthier village, both in economic and development terms, with a Hindu population. Given the preponderance of development and economic class explanations for political behavior and social structure, these are natural criteria upon which to base the diverse case selection. The differences between the two villages allow one to deduce the extent to which the discussion and coordination over kinship networks functions over very different social contexts. At the same time, close observation of the kinship mechanism in these contexts allows the researcher to deduce variation in the strength of the proposed mechanisms.

In many qualitative designs, case studies are chosen carefully from a larger universe of cases; that is, a small number of cases are chosen to deduce causal mechanisms from larger quantitative empirical patterns. In this study, the situation is reversed, the frame for the quantitative empirical analysis is taken to be the the villages under study.[10] There are three justifications for this approach. First, as discussed above, the larger empirical relationship between family as a stated influence is well-established in the Indian context, so there is little need to demonstrate this larger empirical pattern across India. Second, establishing the impact of kinship networks on changes in political opinions and vote choice requires extensive local within village data across family members. Finally, conducting survey research concurrently with qualitative research permits the researcher to bring detailed and focused knowledge of the context through direct observation to explain larger village-level empirical

[10] This is a common method in the study of American political behavior, where cities are often taken as the frame for careful empirical studies (Berelson et al., 1954; Huckfeldt and Sprague, 1995; Gerber and Green, 2000).

patterns.

4.3.2 Political Opinion Formation in Ranjanpur and Chaandinagar

The selected villages are in the Magrahat Purba assembly constituency, which is approximately 70% rural according to the 2011 Indian census. The boundaries of the constituency are coincident with Magrahat 2 block in the district of South 24 Parganas. According to the 2001 Indian census (the latest census for which religious data are available), the constituency is 47% Muslim, well above the state average of 25%. This rural, Muslim character of the constituency largely defines the set of politically salient issues in the area, while Hindu-Muslim tensions are relatively low owing to the unique cultural character of this region in West Bengal.[11]

The area is on a major rail line, and between 30 and 90 minutes south of various points in Kolkata by rail. While still sufficient for basic agricultural production, this particular region does not produce as much as the more fertile lands in other parts of West Bengal. The relative ease of accessing Kolkata, combined with slightly lower agricultural production, creates a larger wage premium for non-agricultural work and significant pressure to engage in day labor or other work connected to Kolkata. As such, villages in Magrahat Purba are reasonably connected to the political demands and information emanating from Kolkata. The two villages, Ranjanpur and Chaandinagar, fall within the same geographical area insofar as they are serviced by the same train station. At the same time, they are approximately a 45 minute walk apart from each other. This distance was selected to minimize spillovers across study villages.

The campaign began with the announcement of candidates from each party. The TMC/Congress alliance selected Namita Saha, a early supporter of Mamata Banerjee, the charismatic leader

[11] Until recently, this area of southern Bengal was heavily forested, as can be deduced from a large shrine to "Bonbibi." As the story goes, Bonbibi was an orphaned girl chosen by Allah to be a 'mediator of peace,' who guaranteed protection of the resources of the forest and all of its citizens, regardless of religion or caste (Jalais, 2010). Today, Bonbibi is still worshipped by Hindus and Muslims alike. However, significant divisions do persist, as can be seen in the non-commensality between Hindus and Muslims, and local political leaders continue to be wary of potential Hindu-Muslim violence.

of TMC.[12] She was a political veteran who was known as somewhat of a political operator, and was widely expected to be selected for the candidate nomination. On the other hand, CPM, in a bit of surprise, selected a very young student leader, Chandan Saha, from the Students' Federation of India (SFI) from a nearby college. The SFI is broadly associated with CPM, and many of CPM's workers and leaders have come through SFI's ranks. Importantly, the candidates were not fully known ahead of time, so their announcement injected new information into the process of forming political preferences.

The political organization of the parties can shed light on how political actors and campaigns affect voter preferences. India's panchayat system is a three-tiered nested system, with the *zilla parishad* (district-level panchayat), *panchayat samiti* (block-level panchayat), and *gram panchayat* (village-level panchayat). Local politics is typically coordinated by block-level party leaders, who are associated with the panchayat samiti. This represents the lowest level at which political actors are relatively professionalized, with dedicated party headquarters that coordinate local party behavior. The panchayat samiti in Magrahat Purba, like many others across India, is housed in the same building as the block development officer (BDO), the lowest-level civil service bureaucrat in charge of executing government policy. Owing to this proximity, partisan responses to administrative decisions are crafted quickly.

At the village level there are two types of party workers, those that are more professionalized and look to organize party matters at the block level and those that deal with matters within the village. Block-level workers are those who can help to organize mass events and carry out the tasks of coordinating village-level party matters. Village-level workers usually work through informal organization, strategizing at tea shops and other meetings spots within the village. In addition to canvassing, they provide the crucial service of "counting" supporters. These counts are based upon direct observation of villagers. On voting day, these village-level workers from each party sit outside polling booths keeping a tally of exactly who enters the booth and the expected vote outcome. In a world where sophisticated microdata

[12] Mamata Banerjee formed the TMC as a breakaway party from Congress in 1997 (formally founding the party in 1998). She is viewed as a strong charismatic leader who led agitations against CPM's land policies in Singur and Nandigram, West Bengal and is currently the Chief Minister of West Bengal.

on voters is unavailable, this "counting" structure provides a flawed, but necessary, substitute as well as a monitoring device for voters.

Ranjanpur

Ranjanpur is subdivided into "paras" or neighborhoods that are titled after the last name of the villagers living in the neighborhood. Since villagers in the same neighborhood share a last name, they are understood to be part of the same extended family. Ranjanpur is a Muslim village and underdeveloped in comparison to many other villages in the area. Most village roads remain unpaved, and the village is often flooded during the monsoons because it sits on particularly low-lying land. The larger structure of political support is conditioned by two major factors: family history and economic wealth.

Priors about political opinions and party support are first formed from the last name of the individual, which is consistent with the name of a particular neighborhood. It is understood that inhabitants of a particular neighborhood are part of the same extended family. Thus, at a very broad level, partisan identity is associated with family. Traditionally, large landowning families, or ex-zamindari families, tend to vote for TMC or Congress due to the losses of land described above at the hands of the CPM.

A second major factor in Ranjanpur's political identity is class. Approximately, three to four generations ago, villagers started specializing in painting buildings and working with plaster across Kolkata. This is still the most common profession in Ranjanpur, but, over time, some individuals have become contractors, becoming significantly more wealthy. Access to contracts typically flows through personal and family networks, and so contractors are clustered by kinship. A second route to greater economic well-being has been government jobs, specifically joining the police force. Government jobs have educational requirements and hiring often works through personal networks. As such, one particular neighborhood has used its kinship connections to bring many family members into the police force. Due to the incentives for education, this is now the most well-educated neighborhood in the village.

There is a class dimension to the politics of CPM and TMC/Congress, and the more well-off families have a tendency to support TMC/Congress. Owing to the extended family culture of Ranjanpur, the leadership of TMC/Congress and CPM are dominated by the two numerically largest extended families in the village. The TMC/Congress-controlling family is broadly more well-off and a former zamindari family, whereas the family that controls CPM still has a significant portion of its family that remains undereducated and involved in day labor.

In short, the structure of political identity in Ranjanpur is intimately tied to kinship. Kinship networks generate economic opportunity and social class, which then structures partisan support.

Chaandinagar

Chaandinagar is a large village, and this study only covers a portion of the village and consists of families in a single polling booth. It is a Hindu area, consisting of a “general caste” neighborhood, and a poorer scheduled caste neighborhood. Unlike Ranjanpur, family sizes are much smaller and many different last names, among those who are seemingly unrelated, can be found in the same neighborhood. In this sense, family is less structurally salient. At the same time, while families are geographically delineated as in Ranjanpur, family identities play a large part in political opinions. Chaandinagar is quite a bit more developed than Ranjanpur, having its own athletic grounds and swimming pool, as well as being located next to a high school. Much like Ranjanpur, political identity is intimately tied to kinship through economic opportunity and social class.

Families in Chaandinagar acquired wealth through two distinct paths. First, the village is home to what is reputed to be a *naib* family. The naib was an individual who managed the lands of a large landowner, and thus inherited a significant share of land. These lands were used for the athletic grounds. Members of this family are typically well-educated, some of them holding upper middle class office jobs in Kolkata.

Second, a large number of families have taken up the skilled labor of silver work. Typically, a subcontractor within the village will act as a middleman carrying goods to and receiving contracts from the Burra Bazar marketplace in Kolkata. While the subcontractor accrues a significant amount of wealth, silversmiths often earn a significant wage as compared to day labor. As silver work is a skilled trade, apprenticeship usually occurs within the family. These wealthy families are clustered within the general caste neighborhood, which, adhering to the class dimension of Bengali politics, tends to vote heavily for TMC and not for CPM. Families in the scheduled caste neighborhood on the other hand rely on other professions, either as day labor or handicraft embroidery of saris, which are far less lucrative, and are more likely to support CPM. The structure of political leadership is a bit more disjointed in Chaandinagar.

All of the major political leaders are associated with the general caste neighborhood, perhaps owing to to the importance of caste in the social structure. Since there are no natural connections for the CPM in the general caste neighborhood, the leadership is made up of family members [and family wings] which broke off from traditionally TMC/Congress-supporting families, in particular the family of the naib. This also demonstrates that when there are party switches, they often involve a particular branch of the kinship network.

4.3.3 Comparing Kinship and Personal Networks in Ranjanpur and Chaandinagar

This paper adopts the kinship network as the structure over which to conduct the analysis. The word "family" is one that makes no claim on structure and social distance, and thus is hard to use in a meaningful analytic way. In Ranjanpur, is everyone in the same neighborhood in the same family, or is it just individuals in the same dwelling, and how does one draw these borders? Virtually any definition of the word "household" is too small a unit for analysis. Two brothers may very well be a part of two different households, but they may still share close kinship relations and engage in political discussion. The kinship network in the analysis accounts for those individuals who may engage in political discussion and coordination with each other due to common kinship, while accounting for the fact that they may come from different households. The kinship network structure also allows for the fact that individuals who are connected within it may differ in social distance (e.g., two women married into the same family are more distant than two brothers). Interestingly, the English word "family" is often used in common parlance in both villages to denote the kinship network as conceived in this paper. This gives some face validity to applying the concept in this setting.

The density and importance of kinship networks, and personal networks more generally, vary quite a bit in Ranjanpur and Chaandinagar, as will also be borne out in the quantitative data. One of the first observable differences in the density of personal networks between the two villages is that any villager in Ranjanpur essentially knows exactly where every other villager in Ranjanpur lives, whereas this is not true in Chaandinagar. The difference in density of kinship and personal networks can be partially understood through differences in marriage practices.

Ranjanpur practices endogamy, or consanguineous marriage, which is common among the Muslim community in India (Bittles, 2002). This is one reason why neighborhoods in Ranjanpur are consistent with the last names of the individuals contained in them. As Ranjanpur is a far poorer village than Chaandinagar, the marriage prospects for men, in an arranged marriage system, are significantly weaker. Even when marriage is not consanguineous, wives tend to come from nearby villages due to the weaker drawing power of men in Ranjanpur in the marriage market. This results in dense but locally concentrated kinship and personal networks in Ranjanpur.

Chaandinagar, by contrast, is both a Hindu village, with lower rates of endogamy, and a more well-off village. The set of marriage partners come from a much a wider base of villages across West Bengal, and sometimes even the city, due to better economic conditions. The resulting personal networks in Chaandinagar are less dense but more spatially dispersed. Spatial variation in kinship networks makes individuals more able to mitigate local consumption shocks (Rosenzweig and Stark, 1989). Furthermore, a broader class of "weak ties" due to spatial dispersion in kinship may allow individuals to access a wider array of economic opportunities (Granovetter, 1973). At the same time, lower kinship network density, combined with higher economic status, in Chaandinagar might make families both less able to enforce coordinated behavior and less dependent upon it.

Family Discussion and Coordination over the Campaign

The qualitative research suggests that there are a number of structural and historical reasons for families to have similar political preferences; these differ quite significantly across the villages of study. At the same time, the role of the family discussion and coordination is common across both Ranjanpur and Chaandinagar. In fact, the existing cooperation across a kinship network required for economic access and social class creates a natural environment for political coordination. At the same time, it is clear that common political preferences due to common histories and those due to persuasion and coordination are analytically distinct. This paper isolates the effect of coordination and discussion across kinship networks on changes in political preferences.

In contrast to urban areas, which were inundated by chaotic political rallies and parades, villages experienced a quieter campaign season. Apart from a few visits from important politicians and the occasional procession through rural areas, the villages were largely isolated from mass political demonstrations. To the extent such activities did occur, they were most often organized near the train station or at a busy market in order to maximize exposure. The chief form of campaigning in the village setting was door-to-door canvassing. Given the heavy hours required for day labor for many villagers, much of this activity would take place at night. Since the canvassers were themselves villagers, the village campaign took on a more personalistic character. An aspect of the political vernacular of the campaign season was the conspicuous use of kinship-based language in political engagements. Political leaders would refer to *ghars* (dwellings) of support, and villagers were open about the types of discussion taking place within the family.

In both of these villages, families take on importance vis-à-vis political identity. Kinship networks, as argued above, occupy a prominent role in structuring economic opportunities and social class for villagers. It is no surprise, therefore, that there is a strong correlation between kinship and political preferences. Preferences are not only a function of shared family histories and common social identity, they are a product of family-level coordination. Families often help mitigate individual-level consumption shocks and engage in resource-sharing, and thus developing a family-level political preference is often desirable. A family that is unable to coordinate its voting behavior is a family that is unable to exert its weight.

This is not to say that individual do not have agency; rather, the codependence among family means that preferences within a kinship network are inextricably linked and coordinated upon. The relationship between kinship networks and political identity seems less to be about a slow political socialization and more about periods of negotiation and coordination within the family.[13] It is only during important moments that kinship relations will take the time to pool information and re-evaluate political positions.

This was explicitly seen during the campaign. Families met to collectively discuss/coordinate vote choice shortly before voting day. A lot of weight is typically accorded to a head of the household in these discussions, but this coordination is complicated since an extended family typically has many heads of households, and the primary breadwinner may not be the patriarch. These meetings offer an opportunity to pool information about the election and strategize over vote choice. Anecdotally, pre-election polling suggesting TMC would easily form government by large media houses (and the discussion around them) had a large impact on decisions about the vote. An election pre-poll conducted jointly by Star-Ananda and Anandabazar Patrika, the largest news channel in West Bengal and the largest newspaper in West Bengal, respectively, predicted the Congress-Trinamool Congress coalition to win 215 out of 294 seats. A second impact was frustration over the land policy and weak economic development under the incumbent CPM. These issues, in addition to explicit incentives for coordination, provided the majority of substance for discussion across kinship networks.

Political leaders explained that their methods of monitoring partisan support, and counting, were based on demonstrated support and that overall support was very difficult to gauge in a secret ballot setting. In particular, leaders mentioned that they could conclusively deter-

[13] The importance of kinship in political identity sometimes causes difficulties for a newly married woman, who must balance between her own family and the family into which she has married; this can be a source of marital friction.

mine a supporter by those who may themselves "close" during the campaign season, through party activism and engaging conversations with other party members.[14] At the same time, it was clear that there was a certain segment of the population that could not be read by the political leaders. These were people who associated with leaders and workers from both parties, and seemingly promised their votes to both of them. This suggested that partisan identity was strategically invested in by families, as opposed to foisted upon them, and that kinship networks provided a space that was relatively immune from pressures above.

4.3.4 Survey Protocol

The population for the survey sample was taken to be the those individuals on the corresponding polling booth's official voter list for the two villages, which is available online from the Elections Commission of India (ECI). An individual is eligible to be registered to vote once he/she reaches the age of 18. Since the voter ID card is the principal form of identification in India, much like a driver's license in the US, essentially all eligible individuals register to vote. The voter list is a good source for family network information as each entry includes a family relationship (usually father or husband), which provides information for a basic family network rendering.[15]

The survey was conducted in two phases, a pre-test and a post-test phase. In India, political parties, media, and researchers are subject to the so-called "model code of conduct." This restricts media and researchers from collecting political data and political parties from making new policy promises. Only campaign behavior is allowed during the model code of conduct, so a pre-post survey that bookends this campaign period provides a good measure of campaign effects. The pre-test took approximately one month and ended the day before beginning of the model code of conduct. The post-test took approximately one month as

[14]Party leaders were open about their engagement in using money during election to buy votes, but even they felt it had little impact due to the secret ballot.

[15]However, these lists are often inaccurate, including names of deceased and people who no longer live in the village (most commonly due to marriage). In India, the voter ID card is generally used as a basic form of identification, much like a driver's license in the United States, and as such, people may hold on to voter ID cards to the village, even if they no longer reside there. The initial phase of the study involved vetting the village for residence.

well, and took place approximately one week after the vote results were announced.

In the pre-test, basic demographic information was collected about each individual, along with a first round of questions on political preferences, including: 1) vote choice, 2) opinions on local issues, 3) opinions on state-level issues, and 4) political demands. Finally, in the first round, data were collected on certain aspects of the individual's social network, such as: 1) friends, 2) preferred tea shop, 3) preferred social club, 4) individual turned to for a loan, and 5) individual turned to when needing to go to the hospital.

In the post-test, questions on the political preferences were repeated. In addition, new network data was collected on: 1) family relations in the village that cannot be gleaned from the voter list (e.g. two sisters married into the same village), 2) participation in women's groups, 3) land contracts between families, and 4) employment contracts between individuals. The data in these paper are drawn from voter preferences in the pre-test and post-test and a family network coding based upon the voter list.

The survey protocol was designed to: 1) derive a sufficient sample to estimate network effects, and 2) elicit truthful responses of private political information.

Villagers in India have very irregular schedules at home due to seasonal employment, day labor, and agricultural priorities, so the surveyor requires a careful strategy to boost response rates. Over the one month period in each phase, the survey team mapped out the schedules of all potential respondents. Surveys were conducted in morning/afternoon and evening shifts, with repeat visits to potential respondents to confirm refusal to participate or non-residence in the village.

The assembly elections were conducted under volatile security conditions which required the stationing of national paramilitary troops during the election. As such, along with a team of 8 surveyors, A coding protocol was created to protect the privacy of each respondent. Each survey was broken into four sections: 1) name sheet, 2) demographic and network information, 3) political preference information, and 4) vote choice. Each section of the survey was identified by a unique code that could only be connected to an individual by the surveyors. In the course of the survey, once the name of the respondent was written on the survey, the name sheet was separated from the rest of survey and kept with the surveyor. Each surveyor carried a large “ballot box.” After the network and preference sections of the survey were completed, they were separated from the survey and dropped into the ballot box. Finally, each respondent was asked to fill out a sample ballot in private, fold up the ballot and drop it in the ballot box. This protocol had the advantage of demonstrating intent to keep information private as well as the fact that, even if our data were seized by others, the information could not be tracked to any individual. This protocol was necessary to elicit truthful responses in a volatile setting that posed potential risks for the respondents.

4.3.5 The Campaign Period

Unlike many other places, the campaign period is well-delineated in India. Campaigns essentially starts with the announcement of candidates and the model code of conduct. The model code of conduct (MCC) promulgated by the Election Commission of India (ECI), a non-partisan constitutional body with wide-ranging powers, helps significantly with this task. The MCC puts strong restriction on the behavior of political actors, media, and researchers during the campaign, which helps dramatically focuses plausible sources of impact over the campaign. The directives under the MCC are followed fairly strictly since behavior is carefully monitored by rival political parties, and the ECI has a high level of independence from political actors.

Under the MCC, government actors can neither announce new policies nor can they process or release new funds under existing welfare and beneficiary schemes. Furthermore, political advertisements in mass media are strictly regulated by the chief electoral officer of the state electoral commission, which works under the aegis of the ECI. Finally, public rallies were effectively banned within 48 hours of the election day. The majority of the impact of the campaign period was restricted to media coverage of campaigns, public rallies and smaller meetings further away from the election date, political deliberation and canvassing nearer to the election date. Finally, local observation by the research team failed to note any serious irregularities during the campaign period.[16]

4.4 Campaign Effects on Vote and Opinion Change

This paper models the influence of kinship network on voter preferences through a pre-post study design over an electoral campaign.[17] The quantity of interest is the average saturated effect of the campaign period, and how it varies over the kinship network. Here, the average saturated effect refers to the average effect under the scenario where each unit in the population experiences the campaign period, inclusive of network spillovers.[18]

The pre-post design, or other longitudinal data, has often been the tool of choice to study the effect of political/electoral campaigns. Two desirable properties for the pre-post design, and their relationship to the estimation of kinship network effects, are discussed in detail here: 1) The ability of pre-post designs to estimate saturated campaign behavior; and 2) the ability to of pre-post designs to capture outcomes at the individual level and remove reverse causality. This section demonstrates that the electoral campaign had an effect on both political opinions and vote shares for the TMC.

4.4.1 Using Pre-Post Designs to Understand Network-Based Campaign Effects

Changes over the Campaign Period

A standard pre-post study design measures the outcome of interest before a specified period (pre-test) and then measures the outcomes of interest again after the period of interest (post-test). Often such designs are structured so that the period includes some "intervention" of

[16] One of the biggest concerns was that the announcement of election results may have had a significant effect upon reported vote choice in the posttest. The estimated effects are in line with other studies such as Patnam (2013). Furthermore, the strict secrecy employed in the survey protocol combined with the concurrent presence of the lead researcher, who was clearly non-partisan, bolstered the quality of the data.

[17] This is also often called a before-after design or a two-stage panel.

[18] Sircar (2014) shows that, in general, randomized experiments cannot retrieve the saturated effect in the presence of spillovers. In particular, under spillovers, the outcome of any unit is dependent upon the treatment status of every other unit. Since a randomized experiment necessarily only treats some subset of the population, the average saturated effect cannot be retrieved from such a design. By contrast, the average saturated effect is retrieved by a randomized experiment when there are no spillovers over the network.

interest. Technically speaking, however, causal attribution in this context can only be given to the entire period between the two measurements, e.g., the campaign period, but not the components, or interventions, within that period, e.g., media exposure, clientelistic appeals (Campbell and Ross, 1968). Thus, we do not typically want to claim that a measurement between two points in time constitutes a "causal" measurement. At the same time, focusing on the measured difference over a period may provide meaningful, interpretable effects.

Brady et al. (2006) make the distinction between potential campaign effects and actual campaign effects. Political campaigns are a function of party workers and leaders making strategic decisions over a portfolio of strategies about how to maximize popular support, as well as strategic decisions by voters on the consumption of various campaign appeals. For instance, party functionaries might believe that it is best to make clientelistic appeals to the impoverished and ideological appeals to professionals. Unfortunately, such decisions are unknown to the researcher, and attempts to directly manipulate a campaign will necessarily fail to account for such decisions.[19]

Potential effects refer to those types of effects that are measured under a controlled scenario that excludes some realistic conditions, such as the personal agency of those creating and those consuming the campaign. These are the types of effects that are measured in randomized control trials and lab experiments, and they are valuable for isolating the effects of a certain intervention, like a message or advertisement during a campaign. In contrast, actual effects refer to those types of effects that do not compromise realistic conditions for the campaign, as in longitudinal studies such as a pre-post design. While the changes in a pre-post design can be attributed to the campaign period, it is typically not possible to deduce the causal effect of individual components of the campaign period because the type and magnitude of campaign exposure are not held constant across the population. In this paper, the phrase "campaign effect" will refer to such pre-post changes, not the causal effect.

In this paper, campaign behavior is envisioned as the equilibrium of strategic behavior of

[19] Although it is common to use the phrase "campaign experiment," such randomized experiments actually manipulate a single piece of information, not entire campaigns which are a mixture of various strategically determined appeals (Wantchekon, 2003).

families and political actors, the sort of effect that cannot be measured with explicit researcher manipulation. In this context, the influence of a kinship network in a pre-post design over the campaign period have an intuitive interpretation–how equilibrium campaign behavior varies within and across kinship networks over time.

Isolating the Influence of Kinship Network

When political parties execute electoral campaigns, they often target families. When an individual is the target of a campaign, it is often likely that another person in her personal or social network will also be targeted. Furthermore, individuals in the same social network share information, discuss politics, and, perhaps, even coordinate voting behavior. Consequently, the net effect of a campaign is much more than being directly targeted by a campaign. In this context, empirically meaningful estimates of campaign effects on any outcome need to account for spillovers and information spreading in a social network, as well as common exposures to the campaign. In this paper, the structure of the kinship network is accounted for using a network autoregressive structure, as detailed below.

Many network studies explicitly deduce claims from correlations over the network, which has been criticized for having poor identification of causal effects (Lyons (2011)). In particular, social relations are often a function of the outcome of interest and vice versa,[20] causing serious endogeneity concerns in the estimates. The most difficult aspect of estimating the effect of a social network upon any outcome of interest is "reverse causality," the fear that the outcome of interest or variables strongly correlated to the outcome of interest will be responsible for the structure of the network. In order to address the concern of reverse causality, this design isolates the effects over a campaign period. One can then investigate how the campaign effects differ across a kinship network that stays fixed over the campaign period. In other words, by limiting inferences to campaign effects, this design isolates the influence an existing kinship structure has upon the outcome of interest.[21]

[20]As an example, similarity in political beliefs between spouses may be due to the fact that spouses discuss politics with each other or because individuals with similar political attitudes tend to marry each other.

[21]It is important to note that these estimated influences are not the same as causal effects. In particular, there is no claim about how manipulating the kinship network affects the outcome of interest. Rather, the

Network-based analyses require an estimate of the effect (of the campaign period) for *each* individual in the network since network heterogeneity occurs at the level of the individual. The difference between the post-test and a lagged pre-test outcome at the level of the individual provides such as estimate.[22] Other common designs, like the rolling cross sections, regression discontinuities, or randomized experiments, provide evidence for average effects at the aggregate level, not the individual level. Accordingly, none of these other methods can easily accommodate the estimation of the effect of spillovers over the entire network.

4.4.2 Campaign Effects on Vote Choice

The data in this paper result from votes collected according to the protocol described in the previous section. The analysis is restricted to individuals who reported casting a vote for either CPM or the TMC/Congress alliance (henceforth, TMC) in order to conduct meaningful before/after analyses with a binary variable. After making these restrictions on the data, there were 837 usable individuals for the analysis in Ranjanpur and 257 usable individuals in Chaandinagar.

In each village, campaign period yields a 10% increase in vote share for TMC. In Ranjanpur, the vote share for TMC jumps from 54% to 64% (from 451 to 535 of 837 voters), and in Chaandinagar the vote share jumps from 68% to 78% (from 175 to 200 of 257 voters). Both of these positive jumps in vote share are highly significant ($p < 0.01$) under the Wilcoxon sign test for paired data. Tables 4.1 and 4.2 display the cross-table of vote shares for CPM and TMC in the pre-campaign and post-campaign phases in the two villages.

To the casual observer, a ten percentage point swing may seem quite high, but "bandwagon effects" are known to be quite strong in India. This a function of the political coordination discussed in section 4.2. For instance, using a geographic discontinuity design and election results, Patnam (2013) finds that surprises in exit poll data yields a twenty per-

kinship network is treated as a "pre-campaign" variable, and the approach detects how particular campaign effects vary across the kinship network. This is a standard technique for isolating the effects of structural or identity-based variables on an outcome of interest, e.g., the effect of gender on the success of a job-training program.

[22] The lagged effect is required because the pre-test outcome may not perfectly predict the post-test outcome.

		Post-Campaign		
		CPM	TMC	
Pre-Campaign	CPM	233	153	**386**
	TMC	69	382	**451**
		302	**535**	

Table 4.1: Ranjanpur Votes

		Post-Campaign		
		CPM	TMC	
Pre-Campaign	CPM	44	38	**82**
	TMC	13	162	**175**
		57	**200**	

Table 4.2: Chaandinagar Votes

centage point increase in support for the winning party. Presumably, the effect is smaller in this sample because there was some awareness among the population that TMC would win the election. Furthermore, the data show that a significant share (especially in Ranjanpur) actually switched their vote to the losing party. The magnitude and direction of the effects, combined with the design, provide strong evidence for believable measurements for the vote choice data. Overall, there is strong evidence of a sizable vote swing towards the winning party (TMC) over the campaign period.

4.4.3 Campaign Effects on Opinion

The opinion data in this paper consists of "ideal points" generated from a 2-parameter Rasch model. The ideal points are generated from the following 7 questions (all as agree/disagree questions) that were asked before and after the campaign. The incumbent government refers to the then outgoing CPM government:

- **P1.** The incumbent government of West Bengal has not attempted to create job for Muslims.
- **P2.** The incumbent government has not been very focused on developing industry.
- **P3.** It was inappropriate for the incumbent government to take land from farmers in Singur and Nandigram.
- **P4.** Mamata Banerjee has a plan for the land in Singur.
- **P5.** The incumbent government has explicitly attempted to take land from Muslims.

- **P6.** It is inappropriate to build the "Salim Rasta."

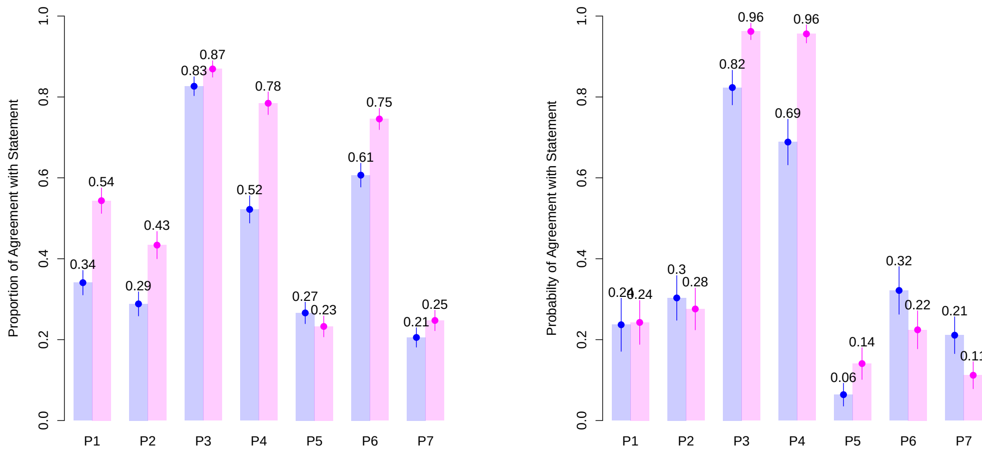
- **P7.** The incumbent (CPM) government hasn't done anything over the last 34 years.

Several points are worth noting about the list of statements above. First, the questions have been transformed from a 4-point scale. Second, the questions listed have been transformed from the original question so that they all have the same orientation (agreement would be consistent with the position of TMC), which is required for the estimation of ideal points. Finally, the questions were chosen to be closely tied to prominent campaign issues over which the CPM and TMC disagreed during the election. The issues were tethered to partisanship for three reasons: a) partisanship is highly salient in West Bengal, b) connection to partisanship makes ideal points over a single dimension more likely, and c) partisan issues allow for assessment of the consistency between issue preference and vote choice. Furthermore, in order to determine campaign effects, it is important to investigate the issues that were explicitly discussed during the campaign.[23] Figure 4.1 displays the overall proportion supporting each of the issues before and after the campaign. In each case, the data were restricted to the voter sample constructed above for those who gave a preference on at least one of the issues, P1-P7, in both pre-campaign and post-campaign phases. This yields 243 respondents in Chaandinagar and 817 in Ranjanpur.

A couple of things are worth noting after looking at figure 4.1. There seems to be a broad movement towards TMC-oriented opinions from the pre-campaign phase to the post-campaign phase. However, there is some variance in the extent of movement, as P6 and P7 actually move in the CPM direction in Chaandinagar, and P5 doesn't move much in Ranjanpur. This suggests that there is some variation in the movement of opinions across separate issues and geography; more importantly, it shows that movement in vote choice doesn't map cleanly on to movement in opinions.

[23]Each of the statements above referred to a major local campaign issue. In particular, the incumbent (CPM) government was criticized for four things: a) poor treatment of Muslims, b) problematic land grab policies for industry (particularly in Singur and Nandigram), c) inability to execute or support of controversial industrial policies, and d) malfeasance during its time in government. Under these guidelines, most of the statements above should be self-explanatory, except for P6. "Salim Rasta" refers to a controversial proposed highway to be built by the Salim Group of Indonesia under the direction of the incumbent government, which required land from the villages under study.

Figure 4.1: Before-After Comparison on Issue Preferences



(a) Ranjanpur (b) Chaandinagar

Figure 4.1 displays the the proportion of respondents supporting each of the of the issues, P1 to P7, before and after the campaign period

Ideal Point Estimation of Opinion

In this paper, a 2-parameter Rasch model was used to estimate ideal points. Many scholars advocate fitting a 3-parameter model, including what is often called a "discrimination parameter," which puts different weights on the salience of the issues (Jackman, 2001). In the 2-parameter model, each issue is given a position on the issue spectrum, but the issues have identical weight in the estimation. In this model, one end of the spectrum will correspond to views consistent with the positions of the TMC, and the other end will be consistent with the positions of the CPM. Unfortunately, the 3-parameter model generally requires strong prior beliefs about the ideological position of each issue which is avoided since there are only 7 questions. The benefit of fitting the 2-parameter model is that it can be fit fairly quickly without strong assumptions on the parameters.

After each issue is placed on the issue spectrum, the model estimates the probability that an individual will agree with the statement (P1 through P7). The higher the probability of agreeing with the statement, the more the individual will be placed towards the TMC side of the issue spectrum. People who agree with positions where very few people agree with the TMC position on the issue will be placed further to the TMC side of the issue spectrum. Finally, in order to estimate the model over two periods, before and after the campaign, underlying position of any given issue does not change over the study period (a fairly reasonable assumption given the short window of the study). Let y_{ik} be the response (agree/disagree) of person $i \in \{1, \ldots, n\}$ on issue $k \in \{1, \ldots, K\}$. The standard 2-parameter Rasch model estimates:

$$P(y_{ik} = 1) = logit^{-1}(\alpha_i - \beta_k) \tag{4.4.1}$$

where α_i denotes the ideal point of person i and β_k denotes the position of issue k on the issue spectrum. Notice, however, that the model is not identified since one can add a constant to α_i and subtract it from β_k. Normally, as is done here, the expected value of α_i is set to 0 to keep the model identified. Now consider issue beliefs in both the pre-campaign and post-campaign phases. Let y_{ikt} denote the value of y_{ik} in period $t \in \{0, 1\}$. There is now a second problem for the analysis. In order to deduce changes in issue beliefs, the changes must occur with respect to the "same" issues. Thus, one must freeze the β_k terms across $t = 0$ and $t = 1$ and estimate separate ideal points, α_{i0} and α_{i1}. To estimate the model, essentially α_{i0} and α_{i1} are treated as ideal points for two separate individuals. However, this form of estimation permits the ability to compare changes from α_{i0} to α_{i1}. The entire 2-parameter Rasch model across the pre and post periods may now be written over a population of n persons in periods $t \in \{0, 1\}$:

$$P(y_{ikt} = 1) = logit^{-1}(\alpha_{it} - \beta_k) \tag{4.4.2}$$

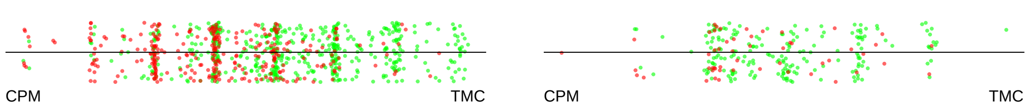
where

$$\alpha_{it} \sim N(0, \sigma_\alpha^2); \quad \beta_k \sim N(\mu_\beta, \sigma_\beta^2)$$

Finally, in order create an interpretable dimension for the analysis, the ideal points (opinions) are formed as $\frac{\alpha_{it}}{\sigma_\alpha}$, where α denotes the entire vector of pre-campaign and post-campaign ideal points. The opinions can be interpreted on a dimension with mean/median 0 and standard deviation 1. Comparing the mean opinion of two subgroups of the populations provides information about relative distance in beliefs between the two groups, where the difference in means can be interpreted in terms of standard deviations over the entire distribution of opinions. The models are fit separately for each village due to difference in salience of the issues (e.g., Muslim issues) across the two villages.

The estimated ideal points in the post-campaign phase are plotted against the vote choice in the post-campaign phase in figure 4.2. The clustering at various points is due to the fact that there are only seven items in the model, and many respondents answer the questions in an identical fashion. As mentioned above, this ideological dimension is expected to be tied to partisan difference, and this is borne out in the figure. The red ideal points denote CPM voters, and the green ideal points denote TMC voters. In Ranjanpur, a CPM voter has a mean ideal point of -0.21, and TMC voter has a mean ideal point of 0.36, so shifting from CPM to TMC yields an increase of 0.57 standard deviations on the ideological scale. In Chaandinagar, the effects are much smaller, where the mean CPM voter has an ideal point of -0.09 and the mean TMC voter has a ideal point of 0.08, suggesting that a switch from CPM to TMC predicts a movement 0.17 standard deviations on the ideological scale. The Mann-Whitney test yields $p < 0.01$ for each of these differences.

Figure 4.2: Post-Campaign Ideal Points and Vote Choice



(a) Ranjanpur (b) Chaandinagar

Figure 4.2 displays the estimated ideal points in the post-camapaign phase on a single dimension, with red points denoting those who voted for CPM and green points denoting those who voted for TMC. There is a strong statistically significant relationship between vote choice and position on the "ideological spectrum" in both villages, suggesting validity for the constructed ideal points. In Ranjanpur in the post-campaign phase, the mean CPM supporter's ideal point is -0.21, and the mean TMC supporter's ideal point is 0.36. In Chaandinagar in the post-campaign phase, the mean CPM supporter's ideal point is -0.09, and the mean TMC supporter's ideal point is 0.08.

A similar pattern is seen in the difference between pre-campaign and post-campaign measurement of opinions in the two villages. In Ranjanpur, the mean ideal point in the population increases from -0.15 in the pre-campaign phase to 0.15 in the post-campaign phase, with $p < 0.01$ according to WIlcoxon sign test with paired data. In Chaandinagar, the mean ideal point in the population increases from -0.05 in the pre-campaign phase to 0.04 in the post-campaign phase, with $p < 0.05$ according to the Wilcoxon sign test with paired data. This suggests that the campaign has strong effects on opinion formation as well. Figure 4.3 summarizes the estimated campaign effects for vote choice and opinions in this section.

	Ranjanpur	**Chaandinagar**
Vote	0.10 (< 0.001)	0.10 (0.002)
Opinion (in SDs)	0.30 (< 0.001)	0.09 (0.031)

Table 4.3: Estimated Campaign Effects for Vote Choice and Opinion by Village

Figure 4.3 displays the differences in estimates for vote choice and ideal points by village for the pre-campaign and post-campaign phases. P-values estimated from a Wilcoxon sign test with paired data are given in parentheses.

This section demonstrates that villages under study experience fairly large shift in vote choice over the campaign period, as well as an associated shift in political opinions. This suggests that ultimately campaigns may have considerable effects on voter behavior, both in vote choice and in ideological opinions, in rural India. The movement of ideological opinions suggests that more than quid pro quo type politics is at play.

4.5 The Influence of Kinship Networks on Vote and Opinion Change

This section investigates the role of kinship networks in the campaign effects deduced in the previous section. In particular, the focus of the section is to deduce an interpretable estimation strategy to understand the changes in vote choice and opinions over the campaign as a function of the kinship network. This section demonstrates that kinship networks have a strong, discernible impact on these changes over the campaign.

4.5.1 Measuring Kinship Networks

The sample population for this study is the set of individuals on the official voter lists of the polling booths corresponding to the villages of study. Voter lists are available online from the Elections Commission of India (ECI). An individual is eligible to be registered to vote once he/she reaches the age of 18. Since the voter ID card is the principal form of identification in India (e.g., which is used for proof of identification for mobile sim card), almost all eligible individuals were registered to vote in the villages studied.[24] The voter list is a good source for the (patriarchal) family network, as each entry includes a family relationship, usually the father for males and unmarried daughters and spouse for women who have married into the village. This provides enough information to generate a family network consisting of spouses, siblings, and parents/children. In this study, a link was formed between two individuals in the kinship network if they were siblings, married, or the parent/child of the other individual. Figure ?? displays an entry from the voter list with identifying information redacted.

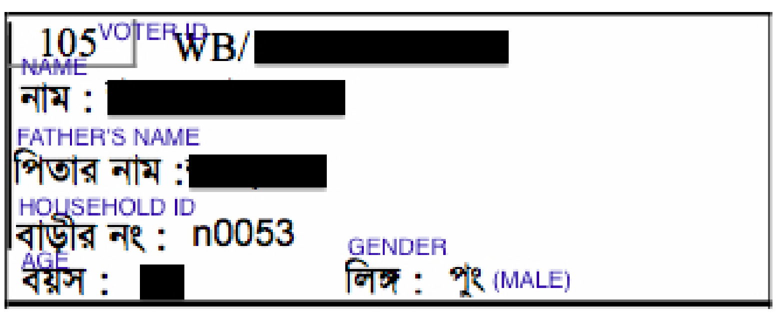


Figure 4.3: Estimated Campaign Effect on TMC Vote Share

Figure ?? shows an example of an entry in the voter list with kinship (and other) information.

In Ranjanpur, there are 731 unique pairs of individuals with a link (dyads) over 837

[24] However, these lists are often inaccurate, including names of deceased and people who no longer live in the village (most commonly due to marriage). In India, the voter ID card is generally used as a basic form of identification, much like a driver's license in the United States, and as such, people may hold on to voter ID cards to the village, even if they no longer reside there. The initial phase of the study involved vetting the village for residence.

individuals satisfying the voting criterion. In Chaandinagar, there are 172 unique pairs of individuals with a link over 257 individuals. The number of links in the network emanating from and individual is typically referred to as the *degree* of the individual. In Ranjanpur, the average degree is 1.75, and, among those individuals with at least one link, the average degree is 2.28. In Chaandinagar, the average degree is 1.34, and, among those individuals with at least one link, the average degree is 1.80. In short, the network sample drawn in Ranjanpur represents more dense kinship relations than in Chaandinagar.

4.5.2 The Relationship between Kinship Networks and Post-Campaign Measures

This subsection demonstrates the existence of an association between kinship and vote choice and political opinions. In figures 4.4 and 4.5, the kinship networks in Ranjanpur and Chaandinagar are displayed by vote choice and ideal points, respectively. In each figure, estimates of Moran's I, a standard measure of "network autocorrelation," is calculated for the post-campaign vote choice and ideal points.

Consider a network characterized by an adjacency matrix, A, such that the entry $A_{ij} = 1$ if there exists a link between i and j, and 0 otherwise. Let W, with entries W_{ij} be the row-standardized weight matrix calculated from A. That is, the terms A_{ij} are divided by the degree of i (if more than 0) so that rows of W sum to 1. In essence, W provides weights over the network to ensure that those individuals with many links do not have disproportionate influence on the constructed measure. For a population of n individuals and outcome y_i for individual i, Moran's I is defined as:

$$I = \frac{n}{\sum_{i \in V} \sum_{j \in V} W_{ij}} \frac{\sum_{i \in V} \sum_{j \in V} W_{ij}(y_i - \overline{y})(y_j - \overline{y})}{\sum_{i \in V}(y_i - \overline{y})^2} \tag{4.5.1}$$

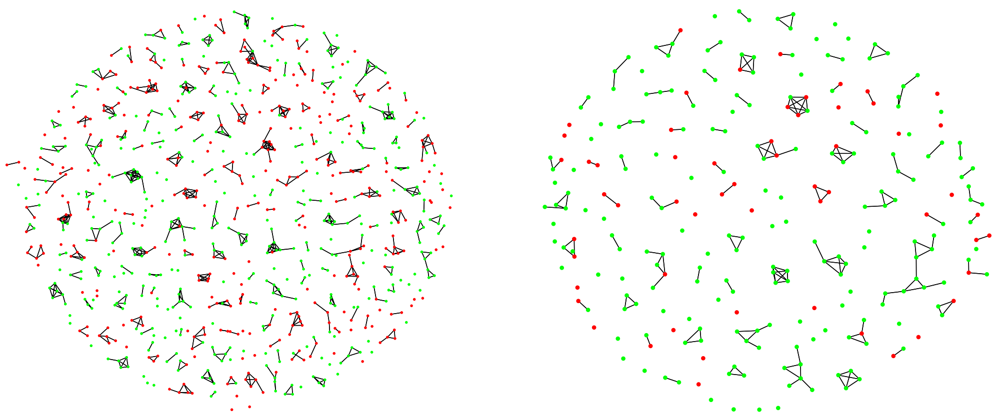
where $\overline{y}$ is the mean of the y_i values.

Moran's I is defined over those individuals who have positive degree (i.e., only over individuals with links). Under these restrictions, the measure is constrained to be between -1 and 1, resulting in its interpretation as a correlation. The estimated Moran's I for figures 4.4

and 4.5 suggest significant network autocorrelation for vote choice and political opinions.

However, this kinship network relationship can be difficult to interpret. It is not clear that the network relationship has anything to do with campaign effects or opinion formation. It may occur due to the fact that those with common kinship start with similar political opinions, as discussed in section 3. In order to separate out the effects for initial vote choice and opinions, the rest of the section describes a technique to deduce vote and opinion change during the campaign over a kinship network.

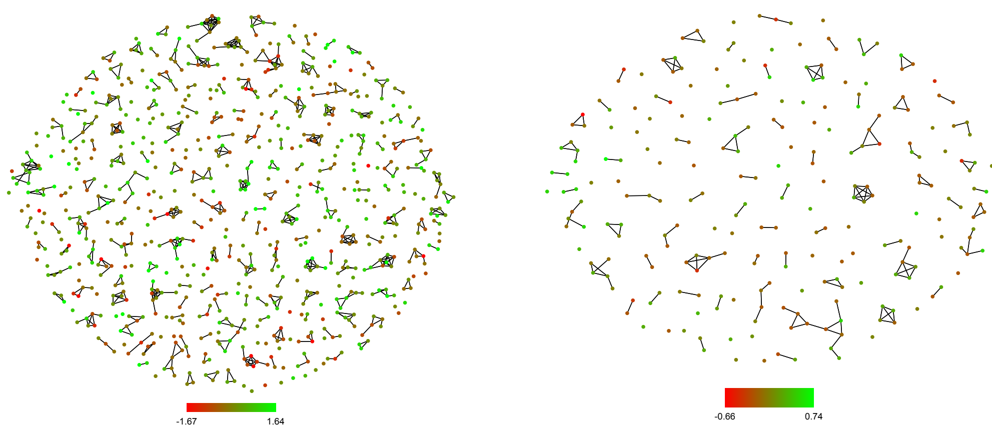
Figure 4.4: Post-Campaign Vote Choice over the Kinship Network



(a) Ranjanpur: $I = 0.33$ (b) Chaandinagar: $I = 0.17$

Figure 4.4 displays the vote choice of respondents overlaid on to the kinship network. In the subfigures, a red vertex denotes an individual who reported voting fro CPM, and a green vertex an individual who reported voting for TMC. In both villages, a significant amount of correlation in behavior is observed over the network.

Figure 4.5: Post-Campaign Ideal Points over the Kinship Network



(a) Ranjanpur: $I = 0.24$ (b) Chaandinagar: $I = 0.18$

Figure 4.5 displays the ideal points of respondents overlaid on to the kinship network. In the subfigures, the color of the vertex denotes (more red or more green) denotes the extent to which the respondent held views more consistent the CPM or TMC positions on the ideological scale. In both villages, a significant amount of correlation in behavior is observed over the network.

4.5.3 A Simple Model of Network Influence and Opinion Change

Consider a population of n individuals arranged over a (kinship) network, $G = (V, E)$, where V (with $|V| = n$) denotes the set of individuals over the network, and $E \subset V \times V$ consists of pairs of individuals that share an undirected link[25] in the network, i.e., direct family ties. Let $y_{it} \in \mathbb{R}$ denote the *opinion* on a particular unidimensional issue for individual $i \in G$ in time period $t \in \{0, 1\}$.

The model presented here describes a general process where individuals who share family ties may influence each other. To develop some intuition, consider the impact of a family member j on individual i and vice versa, that is, $(i, j), (j, i) \in E$. Individuals i and j initially have opinions y_{i0} and y_{j0}, respectively. They engage in a discussion, and reformulate opinions. Between $t = 0$ and $t = 1$, individuals update opinions due to personal characteristics (unrelated to family members), as well as due to the influence of the other family member. When there is no influence of the family link, an individual i updates opinions as a function of

[25] Formally, this implies that if $(i, j) \in E$, then $(j, i) \in E$.

characteristics outside of initial opinion, $\tau_i \in \mathbb{R}$, and relevance of the initial opinion, $\theta_i \in \mathbb{R}$,[26] for future opinion. Therefore, $y_{i1} = \theta_i y_{i0} + \tau_i$. On the other hand if i is fully convinced by opinion of family member j in period 1, then $y_{i1} = y_{j1}$. In reality, however, the influence of a family member is somewhere in between these two extremes:

$$y_{i1} = \gamma_{ij} y_{j1} + (1 - \gamma_{ij})(\theta_i y_{i0} + \tau_i) \tag{4.5.2}$$

$$y_{j1} = \gamma_{ji} y_{i1} + (1 - \gamma_{ji})(\theta_j y_{j0} + \tau_j)$$

$$\gamma_{ij}, \gamma_{ji} \in [0, 1];\ \theta_i, \theta_j, \tau_i, \tau_j \in \mathbb{R}$$

The magnitude of γ_{ij} is a measure of how much influence j has upon i. While this works well for two connected individuals, the analysis requires a method to characterize the expected impact of a family member to an individual over the entire kinship network. Accordingly, the model considers a natural generalization of the process described in equation 4.5.2 to develop a meaningful parameter of interest. For each individual i, j is a family member if it is in the set $N(i)$, the neighborhood of i, i.e., $j \in N(i)$ implies $(i, j), (j, i) \in E$. The cardinality of the neighborhood, $|N(i)| = \delta_i$, is called the degree of i. Once an individual i has many neighbors, one must also consider the relative importance of each family member upon the opinions of i. The captures the fact that j might be quite influential for i in isolation, but when in the context of other family members trying to influence i, j may not carry the importance to influence i heavily in her direction. Let ϕ_{ij} denote the relative importance of j to i. The opinion of i in period 1 can be modeled as the weighted average of influences from her family with weights ϕ:

$$y_{i1} = \sum_{j \in N(i)} \phi_{ij} \gamma_{ij} y_{j1} + \phi_{ij}(1 - \gamma_{ij})(\theta_i y_{i0} + \tau_i); \quad \sum_{j \in N(i)} \phi_{ij} = 1,\ \phi_{ij} \in [0, 1] \tag{4.5.3}$$

Since the goal of the model is to characterize the expected contribution of family member

[26] Intuitively, if the magnitude of θ_i is small, then the initial opinion matters little for future opinion. If, however, θ_i is large and positive, then moving from $t = 0$ to $t = 1$ causes the individual to become more extreme in her opinion.

j to individual i, it will be useful to define three parameters: 1) the relative influence of family member j on individual i – ρ_{ij}; 2) the expected relative influence of a family member on individual i – ρ_i; and 3) the expected relative influence of family members on individuals in the population – ρ. In this analysis, ρ is the parameter of interest. The three parameters are defined formally below:

$$\rho_{ij} = \delta_i \phi_{ij} \gamma_{ij} \tag{4.5.4}$$

$$\rho_i = \frac{1}{\delta_i} \sum_{j \in N(i)} \rho_{ij} \tag{4.5.5}$$

$$\rho = \frac{1}{n} \sum_{i \in V} \rho_i \tag{4.5.6}$$

Each of the parameters defined above is constrained to be in the interval $[0, 1]$. The relative influence of family member j to individual i, ρ_{ij}, has an intuitive interpretation. It is the fraction of the distance j moves i's uninfluenced opinion in period 1, $\theta_i y_{i0} + \tau_i$, towards j's opinion in period 1 (controlling for the relative influence of other family members), and ρ_i is simply the aggregate influence of the family. The parameter of interest, ρ, is simply the average of these aggregate influences from one's direct kinship linkages.

Regression Framework

It can now be show that the parameter of interest ρ may be readily estimated through a network autoregressive regression model. To see this, let $\mathbb{E}_i$ denote the expectation function across individuals, and let $\mathbb{E}_{N(i)}$ denote the expectation across the neighborhood of i. Since opinions in $t = 0$ and $t = 1$ are taken to be observed data, the expectation function is taken conditional upon these values. Taking the conditional expectation, $\mathbb{E}_i[\mathbb{E}_{N(i)}(.)]|y_{0i}, y_{1i}$, on both sides of equation 4.5.3 yields:

$$y_{i1} = \mathbb{E}_i[\mathbb{E}_{N(i)}(\delta_i \phi_{ij} \gamma_{ij})] \frac{1}{\delta_i} \sum_{j \in N(i)} y_{j1} + \mathbb{E}_i[\mathbb{E}_{N(i)}(\theta_i \phi_{ij}(1 - \gamma_{ij}))] y_{i0} + \mathbb{E}_i[\mathbb{E}_{N(i)}((1 - \gamma_{ij})\tau_i)] \tag{4.5.7}$$

Letting $\mathbb{E}_i[\mathbb{E}_{N(i)}((1-\gamma_{ij})\tau_i)] = \alpha$ and $\mathbb{E}_i[\mathbb{E}_{N(i)}(\theta_i \phi_{ij}(1-\gamma_{ij}))] = \beta$ and simplifying yields:

$$y_{i1} = \rho * \frac{1}{\delta_i} \sum_{j \in N(i)} y_{j1} + \beta y_{i0} + \alpha \tag{4.5.8}$$

In matrix form, this equation becomes:

$$\mathbf{y}_1 = \rho \mathbf{W} \mathbf{y}_1 + \beta \mathbf{y}_0 + \alpha \tag{4.5.9}$$

where $\mathbf{W}$ is a matrix with elements w_{ij} such that:

$$w_{ij} = \begin{cases} \frac{1}{\delta_i} & \text{if } j \in N(i) \\ 0 & \text{if } j \notin N(i) \end{cases}$$

The regression form demonstrates a classic endogeneity problem, since the dependent variable $\mathbf{y}_1$ can also be seen on the right side of the equation. Furthermore, the error structure across family members may be very complicated. The trick is to solving these issues is to notice that equation 4.5.9 rewritten by subtracting the first term from both sides:

$$(\mathbf{I} - \rho \mathbf{W})\mathbf{y}_1 = \beta \mathbf{y}_0 + \alpha \tag{4.5.10}$$

where $\mathbf{I}$ is the identity matrix.

One may now run the associated regression (with normally distributed errors) with the transformed dependent variable on the left side with unknown parameters ρ, α, β:

$$\mathbf{y}_1(\mathbf{I} - \rho \mathbf{W}) \sim N(\beta \mathbf{y}_0 + \alpha,\ \sigma^2) \tag{4.5.11}$$

$$\Rightarrow \mathbf{y}_1 \sim N((\mathbf{I} - \rho \mathbf{W})^{-1}(\beta \mathbf{y}_0 + \alpha),\ [(\mathbf{I} - \rho \mathbf{W})'(\mathbf{I} - \rho \mathbf{W})]^{-1}\sigma^2)$$

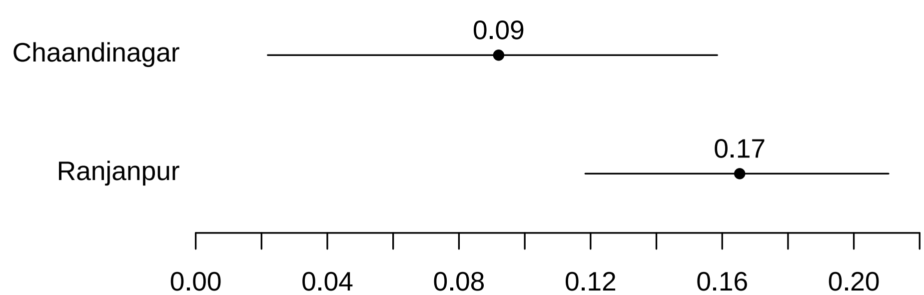
The parameters may be estimated through maximum likelihood estimation. Details on the relative speed and quality of estimation in a maximum likelihood setting for network (spatial) autoregressive regression, vis-a-vis other estimation techniques, may be found in

Franzese and Hays (2008). Causal interpretations of ρ hinge upon the links of the network being independent of underlying individual-level characteristics, which is certainly untrue in most cases.[27] The inclusion of $\mathbf{y}_0$ as a predictor guarantees ρ isolates the expected influence of a family relation on the *change* in opinion; that is, the estimated influence is not due to correlation in initial opinions between family members. Thus, one can interpret ρ as the expected influence of a family relation on changes in opinions in the population over a fixed time period.

4.5.4 Results

The network autoregressive regression model described above was fit to the data in Ranjanpur and Chaandinagar. In particular, the post-campaign vote for TMC and the ideal points estimated post-campaign from the Rasch model were taken as dependent variables, with the pre-campaign vote for TMC and ideal points taken as predictors corresponding to the initial political opinion for the regression form in equation 4.5.11. The models were fit in the `R` statistical environment, using the `lnam` function in the `sna` package. The estimated ρ term from each regression is displayed below.

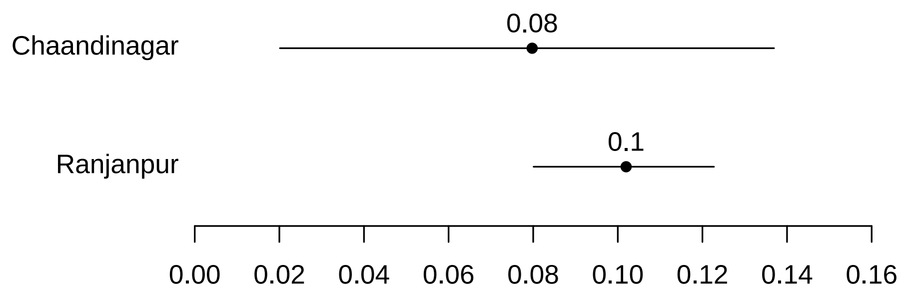
Figure 4.6: Value of ρ for Vote Choice



Figures 4.6 and 4.7 display estimates for the ρ for the vote choice and ideal point re-

[27]In a setting where the links are drawn with probabilities that are not a function of individual characteristics (e.g., the Erdos-Renyi model), the ρ parameter would provide a causal estimate for spillover effect of moving from a null network (no links) to the generated network. This is one general way to deduce the causal impact of the links in a network. Intriguingly, this approach does not require observation of each counterfactual or the randomization probabilities.

Figure 4.7: Value of ρ for Ideal Points



Figures 4.6 and 4.7 display the estimated ρ for vote choice and ideal points by village with 90% confidence bounds simulated from the asymptotic variance-covariance matrix of the estimated parameters (inverse of the Fisher information matrix).

gressions in each of the villages. The estimates are displayed with 90% confidence bounds simulated from the asymptotic variance-covariance matrix for the estimated parameters (using the inverse of the Fisher information matrix). The data suggest a very strong kinship network effect on both vote choice an political opinions. In Chaandinagar, moving from a situation where one's kinship linkages completely support the CPM to a situation where one's linkages completely support the TMC predicts a 9% increase in the probability of voting for TMC, and Ranjanpur displays a stronger effect with such a change predicting a 17% increase in the probability of voting for TMC. By contrast, the models yield changes of similar magnitude with respect to opinions. In Chaandinagar, changing the average ideal point of one's kinship linkages by one standard deviation yields a 0.08 standard deviation movement in ideal points in the same direction; in Ranjanpur, this movement yields a 0.10 standard deviation movement in the same direction. These data suggest strong kinship network effects upon changes in both political opinions and vote choice over the campaign. This demonstrates that kinship networks play an information role in addition to a strategic coordination role for votes, which suggests kinship networks do more than merely engage in material exchange.

4.6 Explaining Kinship Network Effects

This section demonstrates that the observed kinship network effects can be explained by political discussion and coordination within the family. In particular, a majority of respondents report political discussion and coordination within the family for vote choice, and an overwhelming percentage of respondents describe family as the most important influence vis-à-vis other prominent sources of political influence. One of the difficulties in interpreting the impact of kinship networks is that the observed effects may be due to other factors correlated to kinship. This section shows that the results are robust to controlling for other prominent explanations of political influence such as media exposure, associational life, and promises of benefits, as well as relevant demographic factors such as age and gender.

4.6.1 Political Discussion and Coordination within the Family

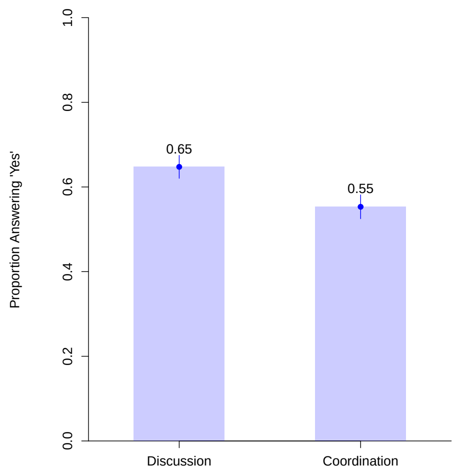
The survey evidence in the villages of study confirms the idea that political discussion and coordination within the value drives observed effect of kinship networks. In the post-campaign phase, respondents were asked the following questions:

- **C1.** Did your family have a discussion regarding the vote (i.e., about vote choice)?
- **C2.** Did your family decide who to vote for together?

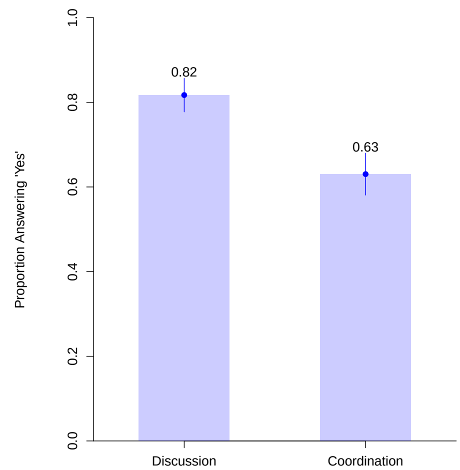
Figure 4.8 displays the proportion of respondents in each of the two villages who reported engaging in political discussion within the family for vote choice (C1) and explicit family-level coordination of vote choice (C2). The vast majority of villagers report engaging in each of these behaviors. In addition to political discussion which may be necessary for information pooling, families tend to engage in explicit coordination of vote choice. These data suggest that families play a crucial role in observable political outcomes. While this speaks to the prevalence of family influence, it does not say anything prominence or importance of family influence.

In order to address the relative importance family influence vis-à-vis other prominent influences on individuals, respondents were asked the most influential information source

Figure 4.8: Proportion of Respondents Who Engage in Family Discussion and Coordination



(a) Ranjanpur



(b) Chaandinagar

Figure 4.8 displays the proportion of villagers in Ranjanpur and Chaandinagar who engage in family discussion and explicit family coordination of vote choice.

for vote choice between family, friends, newspapers, and television news. The results are displayed in tables 4.4 and 4.5.

Source	**Percentage**
Family	83
Friends	4
Newspaper	3
TV News	9

Table 4.4: Ranjanpur

Source	**Percentage**
Family	64
Friends	6
Newspaper	2
TV News	28

Table 4.5: Chaandinagar

In both villages, family is the overwhelmingly prominent source of political influence. The data also point to increasingly important role for television news in political decision-making. Finally, there is evidence that individuals rely more on kinship networks in Ranjanpur and compared the Chaandinagar.

4.6.2 Kinship Network vs. Other Prominent Political Influences

The subsection above shows that respondents attribute the strength of their kinship effects to political discussion and explicit political coordination. Nonetheless, it is possible that the observed effect is due correlation at the family level of other prominent explanations of change in preferences over a campaign. This subsection looks at ρ controlling for prominent sources of influence, namely media, promises of benefits, and associational life, as well as demographic factors of gender and age.[28] The prominent types of influence considered in this analysis are:

- **Media.** As mentioned above, the Columbia School did not believe media effects to be strong, the so-called minimal effects hypothesis, due to the capacity of individuals to select their own personal networks who reinforce their opinions. Since then, there is some concrete evidence of media effects on political opinions, even in the United States (Vavreck, 2001; Gerber et al., 2006; Vavreck, 2009), which has been critical of the Columbia School. Since family members are likely to access similar sources of media, and have similar effects from the media, this may affect the level of kinship effects.

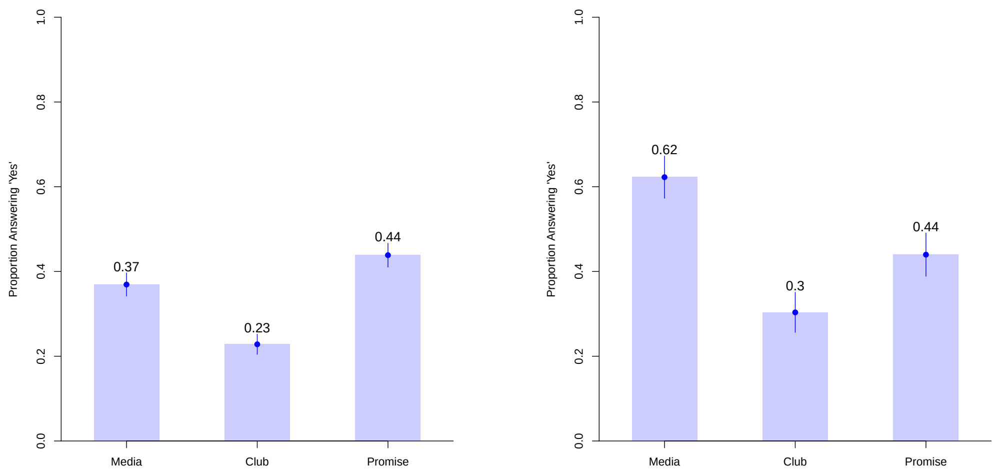
- **Associational Life.** The impact of social capital and "associations" in a robust civil society and democratic behavior has been well-documented (Putnam, 1993). At the same time, Chhibber (2001) has argued that Indian democracy survives with fewer associations among its citizens. To the extent that associations matter in Bengali villages, they are reflected in the social clubs, which are often partisan in nature. Once again, attendance at social clubs is correlated with the kinship network, although it is typically restricted to men.

- **Promises of Benefits.** As described earlier, a major literature focuses on the importance of clientelism and patronage in the Indian system (Chandra, 2004; Kitschelt and Wilkinson, 2007). These promises are expected to be correlated over the kinship network, especially since political actors often target several family members.

[28] Note there is no variation at the household level on social class or identity in this type of data.

The data on media (whether the respondent watches news on television or reads the newspaper) and associational life (whether the respondent attends a social club) were collected in the pre-campaign phase to prevent biases in response. The data on promises were collected from the following question: *"Before the vote, did any party (do not name the party) make promises for personal benefits to you in order to get your vote?"*[29] The relative proportions experiencing each type of influence is displayed in figure 4.9.

Figure 4.9: Proportion of Respondents Exposed to Each Type of Influence



(a) Ranjanpur

(b) Chaandinagar

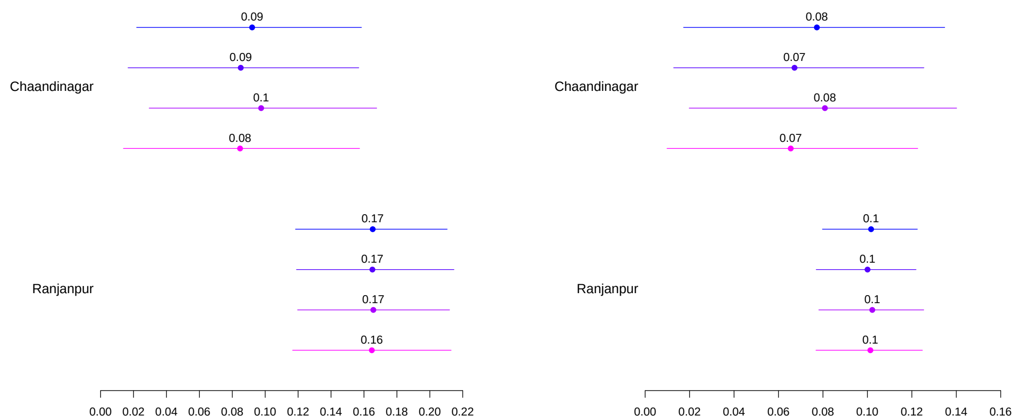
Two categories of predictors were fit to the (saturated) network autoregressive model:

- **Influence.** Pre-campaign ideal point/vote choice, media, associational life, promises (and all higher order interactions)
- **Demographic.** Pre-campaign ideal point/vote choice, gender, age (and all higher order interactions)

Saturated models are fit to purposely overfit the data and provide more conservative estimates of ρ. The results in figure 4.10 show that the value of ρ remains remarkably consistent over all models, suggesting a very robust result.

[29]The explicit instruction to not name a political actor was done to create incentives for truthful reporting.

Figure 4.10: Estimates of ρ for Vote Choice and Opinion Under Various Models



(a) Estimated ρ for Vote Choice (b) Estimated ρ for Ideal Points

Figure 4.1 displays the estimated ρ for vote choice and ideal points by village with 90% confidence bounds simulated from the asymptotic variance-covariance matrix of the estimated parameters (inverse of the Fisher information matrix).

The survey evidence in this section suggests that the influence of kinship networks is due to family-level discussion and political coordination. Even when controlling these for other prominent sources of political influence, one finds similar magnitudes of kinship influence in vote choice and opinion change, suggesting that the lion's share of kinship network influences can be attributed to political discussion and political coordination.

For much of this paper, the data have been presented side-by-side. From a purely quantitative point of view, two cases are insufficient to draw larger trends from the data. However, one can start generating hypotheses from the data presented here alongside qualitative observation. It was noted that the average degree of the individuals sampled in Chaandinagar was lower than that of Ranjanpur. This corresponds to qualitative observation which suggested that social networks in Chaandinagar were less dense than in Ranjanpur.

Chaandinagar broadly does better on all socioeconomic metrics than Ranjanpur. Given the importance of kinship networks, and social networks more generally, for mitigating risk, one might assume the that both network density and the importance of networks for cooperation should be lower in Chaandinagar. While the movement in political opinions are not directly comparable, it does seem that there is little difference in kinship networks on political opinions when measuring in movement according to standard deviations. However, the magnitude of the changes in vote choice is much greater in Ranjanpur than in Chaandinagar. Kinship networks in Ranjanpur play a greater role in the coordination of votes, as compared to Chaandinagar, suggesting variation in the importance of networks in coordination of vote choice. Two related hypotheses result from these observations: 1) the density of personal networks leads to increased reliance on such networks for coordinating voter behavior; and 2) lower reliance on kinship networks for risk mitigation (due to wealth) leads to lower levels of coordination in vote choice over the kinship network.

4.7 Conclusion

Using data from two villages in the Indian state of West Bengal, this paper demonstrates that kinship networks have a strong impact on the formation of political preferences through information pooling of salient issues, political discussion and explicit coordination of the political behavior. Kinship networks affect more than just vote choice, they also affect ideological positions. This suggests that families are engaging in more than just quid pro quo politics with politicians.

A novel approach that juxtaposes qualitative observation at the local-level with micro-level data collection provides strong evidence for the mechanisms proposed. This paper develops an entire empirical strategy to deduce personal network effects on opinion change by integrating longitudinal data over a fixed network with measurement of political opinions through vote choice and ideal point estimation. It is also shown that a network autoregressive model assessing the impact of kinship networks on political preference change can be interpreted from a general decision-theoretic process.

4.7.1 Implications

There are several implications of this study for the future analysis of political behavior in a democratic developing world context. First and foremost, this paper demonstrates that social structures may address concerns associated with low information and weak state capacity in developing societies. This shows how democracy can thrive even in contexts where commonly believed requisites for a robust democracy, such as urbanization, economic development, and high levels of education, are absent. In particular, kinship networks can allow voters to reason through disparate pieces of information in order to make informed choices, while coordination of voting behavior across the kinship network maximizes the impact of the political decision.

A voter who does not have sufficient information, or the capacity, to make reasoned political decisions is vulnerable to manipulation from other political actors. In this way, the kinship network protects a voter and her independence from pressures from above, much like the kinship network helps mitigate consumption shocks in poor, rural settings. While many families may choose to publicly demonstrate their allegiances to a party in exchange for access to state benefits, not all families behave this way. Many other families choose to keep their allegiances private while still coordinating their vote choices. Coordinating vote choice within the kinship network guarantees that the decision has maximum impact on the outcome of the election. In fact, recent work in the Indian state of Rajasthan finds that local political leaders are surprisingly poor at guessing the partisanship of their constituent voters (Schneider, 2014). At the same time, in the same villages, Schneider and Sircar (2014) find that benefits flow through partisanship when such co-partisan affiliation can be inferred by the political leader. This suggests the coexistence of both the clientelistic and non-clientelistic strategies described in this paper. Families can strategically choose to coordinate on a clientelistic strategy by demonstrating their support for a party or a non-clientelistic strategy, and remain insulated from the pressures above. In this sense, the kinship network approach constitutes a more general approach to local political behavior that characterizes a fuller range of voter strategies including information pooling, moving beyond simply clientelism or patronage.

One might be concerned about the generalizability of the results in this paper, given that the study was done in two villages. However, the importance and ubiquity of dense kinship networks in developing rural societies is well established in both anthropology and development economics. Unlike caste relations, which, in their form, is unique to South Asia, and which even varies significantly in practice across Indian states, there is a certain commonality across contexts in using kinship networks to mitigate risk. In that sense, this paper addresses a mechanism that can easily be applied to other settings. However, without explicit examinations in other contexts, it is too early to say the extent to which the results here generalize across the developing world.

4.7.2 Hypothesis Generation

A comparison of the two cases also yields some hypotheses about how the results of this paper might change as economic context and the density of kinship networks change. The economic impact on kinship network effects can be disaggregated into two components, the extent to which the kinship network is dependent upon the village economy and the extent to which individuals are economically dependent upon the kinship network.[30] Ceteris paribus, as individuals become less economically dependent upon kinship or other personal networks, one should expect the effect of such networks to decrease. The implications of economic dependence of the kinship network on the village economy are more complicated. Tables 4.6 and 4.7 display the answer to the following question in the villages of study: *In this election, which level of economic development is most important to you?*

Level	Percentage
National	21
State	13
Village	66

Table 4.6: Ranjanpur–Level for Vote

Level	Percentage
National	33
State	12
Village	55

Table 4.7: Chaandinagar–Level for Vote

Despite being a state-level election, intriguingly very few respondents select the state/province

[30]Unfortunately, these cannot be easily disaggregated in the present study since kinship networks are both less dependent on the village economy and individuals are less hooked into kinship networks in Chaandinagar as compared to Ranjanpur.

level as the most important level of development in the election. This suggests that there are two modes of voters, those who view the state level as making demands to the central government and those that view it as crucial for local development. The percentage of respondents who answer the most local level to this question are a good proxy for the extent to which kinship networks are hooked into the village economy because it characterizes the level at which voters think about sociotropic (societal) issues. While the modal level of importance is the village in both sites of study, a higher percentage of respondents in Ranjanpur are hooked into the village economy, and kinship coordination effects are stronger in Ranjanpur. At the same time, one may conceive of a countervailing force in that typically a voter will have less information about higher levels as compared to the village level. Coordination over a kinship network could result from increased importance of the information pooling function of kinship networks.

One may also wonder what would occur if this study were conducted in an urban setting. If the individual is less dependent upon the larger kinship network, as in the urban middle and upper classes, one might expect voter behavior that is not too different than the Occident; perhaps, kinship networks will have been replaced by friendship networks and online networks. In poor urban areas such as slums, the answer becomes a bit more complex. Short-term migration and questionable tenancy fundamentally fragments kinship networks. In many cases, an urban worker will send remittances to his home village. It is exactly in this setting, without a deep personal network, that individuals are most vulnerable politically. Looking across the literature in highly urbanized countries, as in South America (Auyero, 2001; Stokes, 2005), or even studies of slums in India (Auerbach, 2013) one notices some common features, in particular, the prominence of party machines in co-opting voters and access to the state. A natural hypothesis in this setting is that it is precisely when voters are vulnerable with weaker personal networks that strong party machines are more likely, even in an electorally competitive setting.[31]

[31]While the CPM was dominant in rural West Bengal, it was not an electorally competitive setting. The TMC, which is now electorally dominant, cannot be thought of as a machine party in rural settings due to weak organization.

Chapter 5

Conclusion

In this conclusion, I briefly touch upon my future agenda resulting from the three papers included in this dissertation. Each paper constitutes a distinct line of research that may spawn a number of papers. I wish to reflect on how a research agenda can be created from these three papers.

5.1 On Rent Extraction and Efficient Allocation over Social Networks

This paper creates a method for analyzing the allocation of goods when the units in a social network may extract rents from the allocator. One of the existing challenges in game theory has been the inclusion of social network structures. The "comparative statics" in this universe are no longer a function of parameters, they are a function of network structure. As such, many of the propositions address the effect of changing the structure, e.g., by adding a link.

There are two directions for future research in this game theoretic setting. The first direction is to keep the basic structure of the model but to change the objectives of the allocator. In the model presented in this dissertation, the allocator is required to target every unit in the network. One may imagine that an allocator prefers to target as many units as possible within a fixed budget constraint or a fixed number of units. Changing the objective function of the allocator is likely to create very different results.

A second direction for future research is to bring data to the problem. The natural place to conduct such an analysis is in the provision of local public goods that have spillovers, like schools. This would entail mapping out where spillovers from local public goods flow, and the relative bargaining power of various localities in bargaining with the government.

5.2 Analyzing Randomized Experiments with Spillovers

The paper included in this dissertation is intended to be used as a guiding document for future papers and research design. The paper shows that one can find non-parametric, causally-identified estimates of quantities of interest when one considers the "social distance" between the units of study. As of right now the project is heavy on abstraction and weaker on explicit application of the ideas.

Two major agendas that will come out of this paper are: 1) clarifying the concept of social distance; and 2) applying the method to data with spatial and network spillovers. While the concept of social distance is formally defined in the paper, much more work needs to be done on when the notion of social distance is practically usable. For instance, in spillovers over online networks, it is natural to believe network distance will suffice as a social distance. For any outcome that requires face-to-face contact, such as certain diseases, it is natural to believe that geographic distance will suffice as a social distance. It will also be necessary to develop practical heuristics for when such a candidate for social distance will not suffice for the analysis.

The second agenda is to apply the method to existing data and apply it to new research designs. As described in the paper, thinking about social distance in the design phase is likely to yield major benefits in estimating causal effects after the study has been concluded. The goals here are simple. The method will be applied to existing datasets to see whether standards of causal identification were truly achieved, and to re-estimate effects. Furthermore, the goal is to undertake a series of experiments which explicitly incorporate social distance into the research design to demonstrate applicability of the method.

5.3 A Tale of Two Villages: Kinship Networks and Preference Formation in Rural India

This paper combined qualitative observation with a before-after study design to draw out mechanisms in political preference formation in rural India. The study speaks to my interests in local political behavior in India. In particular, this paper shows how local kinship networks process political information and coordinate political behavior. The data demands for networks are often very high, and the world of "big data" such as online networks, is often too large and confusing to draw out clean mechanisms. The techniques used in this paper may be applied in a wide range of settings to deduce and understand network effects.

Future goals of the study include expanding the logic of the paper to urban areas and state actors. While the analysis has been conducted in a rural setting, it would be advisable to conduct a similar study in an urban setting in West Bengal where kinship networks are likely more fragmented, but other personal networks may dominate social relations. The natural next step in this study is to replicate in such an urban setting.

The paper made a number of claims about how voters interact with the state. However, the state is missing in the data analysis. Along with Mark Schneider, I have developed a set of techniques to cross-reference voter preferences and distributional preferences of local political leaders. The goal is to apply similar methods in rural and urban settings in West Bengal. Finally, it would be advisable to test these theories in other states in India, and perhaps outside of India.

Bibliography

Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Review of Modern Physics*, 74.

Anderson, L. (1987). The state in the middle east and north africa. *Comparative Politics*, 1-18.

Aronow, P. M. and Samii, C. (2013). Estimating average causal effects under general interference. Unpublished.

Auerbach, A. (2013). Demanding development: Democracy, community governance, and public goods provision in india's urban slums. Unpublished dissertation.

Aumann, R. (1959). Acceptable points in general cooperative n-person games. In *Contributions to the Theory of Games IV*. Princeton University Press, Princeton, NJ.

Auyero, J. (2001). *Poor People's Politics*. Duke University Press.

Ballester, C., Calvo-Armengol, A., and Zenou, Y. (2006). Who's who in networks. wanted: The key player. *Econometrica*, 74(5):403–417.

Banerjee, M. (2011). Elections as communitas. *Social Research*, 78(1):75–98.

Bardhan, P., Mitra, S., Mookherjee, D., and Sarkar, A. (2009). Local democracy and clientelism: Implications for political stability in rural west bengal. *Economic and Political Weekly*, 44(9):46–58.

Bardhan, P., Mitra, S., Mookherjee, D., and Sarkar, A. (2011). Political participation, clientelism, and targeting of local government programs: Analysis of survey results from rural west bengal, india. CSSS Working Paper, No. 19.

Bardhan, P. and Mookherjee, D. (2003). Political economy of land reforms in west bengal 1978-98. Unpublished.

Basu, D. (1971). An essay on the logical foundations of survey sampling, part 1. In Godambe, V. P. and Sprott, D. A., editors, *Foundations of Statistical Inference*, pages 203–242. Holt, Reinhart, and Winston, Toronto.

Bates, R. H. (1981). *Markets and States in Tropical Africa: The Political Basis of Agricultural Policies*. University of California Press.

Berelson, B., Lazarsfeld, P., and McPhee, W. (1954). *Voting: A Study of Opinion Formation in a Presidential Campaign*. University of Chicago Press.

Berlin, I. (2002). Two concepts of liberty. In *Liberty: Incorporating Four Essays on Liberty*, pages 166–217. Oxford University Press.

Bernheim, B. D., Peleg, B., and Whinston, M. D. (1987). Coalition-proof nash equilibria i. concepts. *Journal of Economic Theory*, 42(1):13–29.

Bienenstock, E. J. and Bonacich, P. (1993). Game-theory models for exchange networks: Experimental results. *Sociological Perspectives*, 36(2):117–135.

Bittles, A. H. (2002). Endogamy, consanguinity and community genetics. *Journal of Genetics*, 81(3):91–98.

Bonacich, P. (1987). Power and centrality: A family of measures. *The American Journal of Sociology*, 92(5):1170–1182.

Bowers, J., Frederickson, M. M., and Panagopoulos, C. (2013). Reasoning about interference between units: A general framework. *Political Analysis*, 21(1):97–124.

Brady, H. E., Johnston, R., and Sides, J. (2006). The study of political campaigns. In *Capturing Campaign Effects*, pages 1–26. University of Michigan Press, Ann Arbor.

Bramoullé, Y. and Kranton, R. (2007). Public goods in networks. *Journal of Economic Theory*, 135(1):478–494.

Bussell, J. (2014). Representation between the votes: Informal citizen-state relations in india. Unpublished.

Campbell, A., Converse, P. E., Miller, W. E., and Stokes, D. E. (1960). *The American Voter*. University of Chicago Press.

Campbell, D. T. and Ross, H. L. (1968). The connecticut crackdown on speeding: Time-series data in quasi-experimental analysis. *Law and Society Review*, 3(1):33–54.

Census of India (2011). Accessed June 9, 2014.

Chandra, K. (2004). *Why Ethnic Parties Succeed: Patronage and Ethnic Headcounts in India*. Cambridge University Press.

Chandra, K. (2009). Why voters in patronage democracies split their tickets: Strategic voting for ethnic parties. *Electoral Studies*, 28:21–32.

Chen, J., Humphreys, M., and Modi, V. (2010). Technology diffusion and social networks: Evidence from a field experiment in uganda. Unpublished.

Chhibber, P. and Nooruddin, I. (2007). Unstable politics: Fiscal space and electoral volatility in the indian states. *Comparative Political Studies*, 41:1069–1081.

Chhibber, P. K. (2001). *Democracy without Associations: Transformation of the Party System and Social Cleavages in India*. Michigan University Press, Ann Arbor.

Coppock, A. and Sircar, N. (2014). An experimental approach to causal identification of spillover effects under general interference. Working Paper.

Cox, D. R. (1958). *Planning of Experiments*. Wiley, New York.

Crainiceanu, C. M., Ruppert, D., Claeskens, G., and Wand, M. P. (2005a). Exact likelihood ratio tests for penalized splines. *Biometrika*, 92(1).

Crainiceanu, C. M., Ruppert, D., and Wand, M. P. (2005b). Bayesian analysis for penalized spline regression using winbugs. *Journal of Statistical Software*, 14(14):1–24.

Daskin, M. S. (1995). *Network and Discrete Location: Models, Algorithms, and Applications*. Wiley-Interscience.

Downs, A. (1957). *An Economic Theory of Democracy*. Harper.

Duflo, E. and Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics*, 118(3):815–842.

Franzese, Jr., R. J. and Hays, J. C. (2008). Empirical models of spatial interdependence. In Box-Steffensmeier, J., Brady, H. E., and Collier, D., editors, *Oxford Handbook of Political Methodology*. Oxford University Press.

Galeotti, A. and Goyal, S. (2009). Influencing the influencers: a theory of strategic diffusion. *RAND Journal of Economics*, 40(3):509–532.

Gelman, A. and Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, New York.

Gelman, A. and King, G. (1993). Why are american election candidate polls so variable when votes are so predictable? *British Journal of Political Science*, 23(4).

Gerber, A. S. and Green, D. P. (2000). The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment. *American Political Science Review*, 94(3):653–663.

Gerber, A. S. and Green, D. P. (2012). *Field Experiments: Design, Analysis, and Interpretation*. W. W. Norton and Company.

Gerber, A. S., Karlan, D., and Bergan, D. (2006). Does the media matter? a field experiment measuring the effect of newspapers on voting behavior and political opinions. Unpublished.

Giné, X. and Mansuri, G. (2011). Together we will: Experimental evidence on female voting behavior in pakistan. Unpublished.

Glass, J., Bengtson, V. L., and Dunham, C. C. (1986). Attitude similarity in three generational families: Socialization, status inheritance, or reciprocal influence? *American Sociological Review*, 51:685–698.

Goyal, S. (2003). Learning in networks: a survey. In Demange, G. and Woooders, M., editors, *Group formation in economics: networks, clubs, and coalitions*, pages 122–162. Cambridge University Press.

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380.

Hájek, J. (1964). Asymptotic theory of rejective sampling with varying probabilities from a finite population. *Annal of Mathematical Statistics*.

Halloran, M. E. and Struchiner, C. J. (1991). Study designs for dependent happenings. *Epidemiology*, 2(5):331–338.

Haynes, T. W., Hedetniemi, S. T., and Slater, P. J. (1998). *Fundamentals of Domination in Graphs*. Marcel Dekker.

Homans, G. C. (1958). Social behavior as exchange. *American Political Science Review*, 63(6):597–606.

Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*.

Huckfeldt, R. and Sprague, J. (1995). *Citizens, Politics, and Social Communication: Information and Influence in an Election Campaign*. Cambridge University Press, New York.

Hudgens, M. G. and Halloran, M. E. (2008). Toward causal inference with interference. *Journal of the American Statistical Association*, 103(482):832–843.

Huntington, S. P. (1991). Democracy's third wave. *Journal of Democracy*, 2(2).

Inden, R. B. and Nicholas, R. W. (1977). *Kinship in Bengali Culture*. University of Chicago Press.

Jackman, S. (2001). Multidimensional analysis of roll call data via bayesian simulation: Identification, estimation, inference and model checking. *Political Analysis*, 9(3):227–241.

Jaffrelot, C. and Kumar, S., editors (2009). *Rise of the Plebeians? The Changing Face of the Indian Legislative Assemblies*. Routledge India.

Jalais, A. (2010). *Forest of Tigers: People, Politics and Environment in the Sunderbans*. Routledge, New Delhi.

Jennings, M. K., Stoker, L., and Bowers, J. (2009). Politics across generations: Family transmission reexamined. *Journal of Politics*, 71(3):782–799.

Kitschelt, H. and Wilkinson, S. (2007). Citizen–politician linkages: An introduction. In Kitschelt, H. and Wilkinson, S., editors, *Patrons, Clients, and Policies*. Cambridge University Press.

Kothari, R. (1964). The congress 'system' in india. *Asian Survey*, 12(4):1161–1173.

Krishna, A. (2002). *Active Social Capital*. Columbia University Press.

Kruks-Wisner, G. K. (2011). Making claims: Citizenship and service delivery in rural india. MIT Political Science Department Research Paper No. 2011-16.

Lazarsfeld, P., Berelson, B., and Gaudet, H. (1944). *The People's Choice: How a Voter Makes Up His Mind in a Presidential Campaign*. Columbia University Press.

Lewis, O. (1954). *Group Dynamics in a North-Indian Village: A Study in Factions*. Programme Evaluation Organisation, Planning Commission.

Lijphart, A. (2006). *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. Yale University Press, New Haven.

Linden, L. (2004). Are incumbents really advantaged? the preference for non-incumbents in india. Job Market Paper.

Little, R. J. (2006). Calibrated bayes: A bayes/frequentist roadmap. *The American Statistician*, 60(3):213–223.

Little, R. J. (2011). Calibrated bayes, for statistics in general, and missing data in particular. *Statistical Science*, 26(2):162–174.

Lokniti (2009). National election survey. `http://www.lokniti.org/pdfs_dataunit/questionnaires_2009/nes-2009-post-poll-survey-finding.pdf`.

Lucas, R. E. B. and Stark, O. (1985). Motivations to Remit: Evidence from Botswana. *Journal of Political Economy*, 93(5):901–918.

Lupia, A. and McCubbins, M. (1998). *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* Cambridge University Press, New York.

Lust, E. (2009). Competitive clientelism in the middle east. *Journal of Democracy*, 20(3):122–135.

Lyons, R. (2011). The spread of evidence-poor medicine via flawed social-network analysis. *Statistics, Politics, and Policy,*, 2(1).

Mallick, R. (1993). *Development Policy of a Communist Government: West Bengal Since 1977*. Cambridge University Press.

Manor, J. (2000). Small-time political fixers in india's states: "towel over armpit". *Asian Survey*, 40(5):816–835.

Markovsky, B., Willer, D., and Patton, T. (1988). Power relations in exchange networks. *American Sociological Review*, 53(2):220–236.

Michelutti, L. (2009). *The Vernacularisation of Democracy: Politics, Caste and Religion in India*. Routledge India.

Migdal, J. S. (1988). *Strong Societies and Weak States: State-Society Relations and State Capabilities in the Third World*. Princeton University Press, Princeton, NJ.

Parsons, T. (1951). *The Social System*. The Free Press, New York.

Patnam, M. (2013). Learnign from exit polls in sequential elections: Evidence from a policy experiment in india. Unpublished.

Popkin, S. (1994). *The Reasoning Voter: Communication and persuasion in presidential campaigns*. University of Chicago Press, Chicago.

Posner, D. (2005). *Institutions and Ethnic Politics in Africa*. Cambridge University Press.

Przeworski, A., Alvarez, M. E., Cheibub, J. A., and Limongi, F. (2000). *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990*. Cambridge University Press, New York.

Putnam, R. D. (1993). *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press, Princeton.

Rosenbaum, P. R. (2007). Interference between units in randomized experiments. *Journal of the American Statistical Association*, 102(477):191–200.

Rosenzweig, M. R. (1988). Risk, Implicit Contracts and the Family in Rural Areas of Low-Income Countries. *The Economic Journal*, 98(393):1148–1170.

Rosenzweig, M. R. and Stark, O. (1989). Consumption Smoothing, Migration, and Marriage: Evidence from Rural India. *Journal of Political Economy*, 97(4):905–926.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 74(5):688–701.

Rubin, D. B. (1978). Bayesian inference for causal effects: the role of randomization. *Annals of Statistics*, 6:34–58.

Rubin, D. B. (1980). Discussion of "randomization analysis of experimental data in the fisher randomization test. *Journal of the American Statistical Association*, 75:591–593.

Rubin, D. B. (1990). Comment: Neyman (1923) and causal inference in experiments and observational studies. *Statistical Science*, 5(4):472–480.

Ruppert, D., Wand, M. P., and Carroll, R. J. (2003). *Semiparametric Regression*. Cambridge University Press, Cambridge, UK.

Ruud, A. E. (2003). *The Poetics of Village Politics: The Making of West Bengal's Rural Communism*. Oxford University Press, Delhi.

Schneider, M. (2014). Does clientelism work? a test of guessability in india. Unpublished.

Schneider, M. and Sircar, N. (2014). Whose side are you on? identifying distributive preferences of local politicians in india. Unpublished.

Scott, J. (1972). Patron-client politics and political change in southeast asia. *The American Political Science Review*, 66(1):91–113.

Seawright, J. and Gerring, J. (2008). Case selection techniques in case study research: A menu of qualitative and quantitative options. *Political Science Quarterly*, 61(2):294–308.

Sen, A. (1970). The impossibility of a paretian liberal. *Journal of Political Economy*, 78(1):152–157.

Sinclair, B., McConnell, M., and Green, D. P. (2012). Detecting spillover effects: Design and analysis of multi-level experiments. *American Journal of Political Science*, 56(4):1055–1069.

Sircar, N. (2014). Analyzing randomized experiments with spillovers. Unpublished.

Sobel, M. E. (2006). What do randomized studies of housing mobility demonstrate?: Causal inference in the face of interference. *Journal of the American Statistical Association*, 101(476):1398–1407.

Stepan, A., Linz, J. J., and Yadav, Y. (2011). *Crafting State-Nations: India and Other Multinational Democracies*. John Hopkins University Press, Baltimore, Maryland.

Stepan, A. C. and Linz, J. J. (1996). *Problems of Democratic Transition and Consolidation: Southern Europe, South America, and Post Communist Europe*. John Hopkins University Press, Baltimore.

Stokes, S. C. (2005). Perverse accountability: A formal model of machine politics with evidence from argentina. *American Political Science Review*, 99(3):315–325.

Tchetgen Tchetgen, E. J. and VanderWeele, T. J. (2010). On causal inference in the presence of interference. *Statistical Methods in Medical Research*, 21(1):55–75.

Vaishnav, M. (2012a). The merits of money and "muscle": Essays on criminality, elections and democracy in india. Unpublished.

Vaishnav, M. (2012b). Resizing the state. *Caravan*, October 1.

Vavreck, L. (2001). Voter uncertainty and candidate contact: New influences on voting behavior. In Hart, R. and Shaw, D., editors, *New Agendas: Communication and U.S. Elections*. Rowman and Littlefield, New York.

Vavreck, L. (2009). *The Message Matters: The Economy and Presidential Campaigns*. Princeton University Press.

Wantchekon, L. (2003). Clientelism and voting behavior: Evidence from a field experiment in benin. *World Politics*, 55:399–422.

Weiner, M. (1967). *Party Building in a New Nation: The Indian National Congress*. University of Chicago Press.

Wilson, W. A. (1931). On quasi-metric spaces. *American Journal of Mathematics*, 53(3):675–684.

Wood, S. N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(1):95–114.